

5808 # 58

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :

C12N 9/16, 15/55, A23K 1/165

A1

(11) International Publication Number:

WO 99/49022

(43) International Publication Date: 30 September 1999 (30.09.99)

(21) International Application Number: PCT/DK99/00153

(22) International Filing Date: 22 March 1999 (22.03.99)

(30) Priority Data:

0407/98	23 March 1998 (23.03.98)	DK
PA 1998 00806	19 June 1998 (19.06.98)	DK
PA 1998 01176	18 September 1998 (18.09.98)	DK
PA 1999 00091	22 January 1999 (22.01.99)	DK

(71) Applicant: NOVO NORDISK A/S [DK/DK]; Corporate Patents, Novo Allé, DK-2800 Bagsværd (DK).

(72) Inventor: SVENDSEN, Allan; Novo Nordisk A/S, Novo Allé, DK-2880 Bagsværd (DK).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.
With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.

(54) Title: PHYTASE VARIANTS

(57) Abstract

Phytase variants, their preparation and uses, which phytase variants, when aligned according to Fig. 1, are amended as compared to a model phytase in at least one of a number of positions. Preferred model phytases are basidiomycete and ascomycete phytases, such as Peniophora phytase and Aspergillus phytases. Preferred phytase variants exhibits amended activity characteristics, such as improved specific activity and/or improved thermostability.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LK	Sri Lanka	SD	Sudan		
DK	Denmark	LR	Liberia	SE	Sweden		
EE	Estonia			SG	Singapore		

Phytase variants

FIELD OF THE INVENTION

This invention relates to variants of phytases, in particular variants of ascomycete phytases and variants of basidiomycete phytases, the corresponding cloned DNA sequences, a method of producing such phytase variants, and the use thereof for a number of industrial applications.

BACKGROUND OF THE INVENTION

10 Phytic acid or myo-inositol 1,2,3,4,5,6-hexakis dihydrogen phosphate (or for short myo-inositol hexakisphosphate) is the primary source of inositol and the primary storage form of phosphate in plant seeds. Phytin is a mixed potassium, magnesium and calcium salt of inositol.

15 The phosphate moieties of phytic acid chelates divalent and trivalent cations such as metal ions, i.a. the nutritionally essential ions of calcium, iron, zinc and magnesium as well as the trace minerals manganese, copper and molybdenum.

 Phytic acid and its salts, phytates, are often not
20 metabolized, i.e. neither the phosphorous thereof, nor the chelated metal ions are nutritionally available.

 Accordingly, food and feed preparations need to be supplemented with inorganic phosphate and often also the nutritionally essential ions such as iron and calcium, must be
25 supplemented.

 Still further, the phytate phosphorus passes through the gastrointestinal tract of such animals and is excreted with the manure, resulting in an undesirable phosphate pollution of the environment resulting e.g. in eutrophication of the water
30 environment and extensive growth of algae.

Phytic acid or phytates, said terms being, unless otherwise indicated, in the present context used synonymously or at random, are degradable by phytases.

The production of phytases by plants as well as by 5 microorganisms has been reported. Amongst the microorganisms, phytase producing bacteria as well as phytase producing fungi are known.

There are several descriptions of phytase producing filamentous fungi belonging to the fungal phylum of Ascomycota 10 (ascomycetes). In particular, there are several references to phytase producing ascomycetes of the *Aspergillus* genus such as *Aspergillus terreus* (Yamada et al., 1986, Agric. Biol. Chem. 322:1275-1282). Also, the cloning and expression of the phytase gene from *Aspergillus niger* var. *awamori* has been described 15 (Piddington et al., 1993, Gene 133:55-62). EP 0420358 describes the cloning and expression of a phytase of *Aspergillus ficuum* (*niger*). EP 0684313 describes the cloning and expression of phytases of the ascomycetes *Aspergillus niger*, *Myceliophthora thermophila*, *Aspergillus terreus*. Still further, some partial 20 sequences of phytases of *Aspergillus nidulans*, *Talaromyces thermophilus*, *Aspergillus fumigatus* and another strain of *Aspergillus terreus* are given.

The cloning and expression of a phytase of *Thermomyces lanuginosus* is described in WO 97/35017.

25 There is a current need for phytases of amended properties or characteristics, e.g. phytases of increased thermostability, altered pH optimum (a high pH optimum being desirable for in-vitro processing, a low for in-vivo processing in the gastrointestinal tract), and/or of a higher specific activity.

SUMMARY OF THE INVENTION

In a first aspect, the invention provides phytase variants, the characteristics of which are amended - as compared to a so-called model phytase.

5 Any model phytase, which is of a certain similarity to thirteen herein specifically disclosed model phytases, can be made the model of such variants.

In another aspect, the invention relates to a novel phytase derived from *Cladorrhinum foecundissimum*.

10 In still another aspect, the invention provides DNA sequences encoding these phytase variants and this phytase, and methods of their production.

Finally, the invention also relates generally to the use of the phytase and the phytase variants for liberating
15 phosphorous from any phytase substrate, in particular inorganic phosphate from phytate or phytic acid.

BRIEF DESCRIPTION OF THE DRAWINGS

In the detailed description of the invention below,
20 reference is made to the drawings, of which

Fig. 1 is an alignment of thirteen specific phytase sequences (a multiple sequence alignment according to the program PileUp; GapWeight: 3.000;
25 GapLengthWeight: 0.100);

Fig. 2 this figure shows the amino acid and DNA sequence of a first phytase ("P_involutus-A1") derived from strain CBS 100231 of *Paxillus involutus* which was deposited on 28.11.97; the expression plasmid pYES
30 2.0 comprising the full length cDNA sequence was

transformed into *E. coli* strain DSM 11842 which was deposited on 12.11.97 (see WO 98/28409);

Fig. 3

5

this figure shows the amino acid and DNA sequence of a second phytase ("P_involutus-A2") derived from strain CBS 100231 of *Paxillus involutus* which was deposited on 28.11.97; the expression plasmid pYES 2.0 comprising the full length cDNA sequence was transformed into *E. coli* strain DSM 11843 which was deposited on 12.11.97 (see WO 98/28409);

10

Fig. 4

15

this figure shows the amino acid and DNA sequence of a phytase ("T_pubescens") derived from strain CBS 100232 of *Trametes pubescens*, which was deposited on 28.11.97; the expression plasmid pYES 2.0 comprising the full length cDNA sequence was transformed into *E. coli* strain DSM 11844 which was deposited on 12.11.97 (see WO 98/28409);

20 Fig. 5

25

this figure shows the amino acid and DNA sequence of a phytase ("A_pediades") derived from strain CBS 900.96 of *Agrocybe pediades* deposited on 04.12.96; the expression plasmid pYES 2.0 comprising the full length cDNA sequence was transformed into *E. coli* strain DSM 11313 which was deposited on 02.12.96 (see WO 98/28409);

Fig. 6

30

this figure shows the amino acid and DNA sequence of a phytase ("P_lycii") derived from strain CBS 686.96 of *Peniophora lycii* which was deposited on 04.12.96; the expression plasmid pYES 2.0 comprising the full

length cDNA sequence was transformed into *E. coli* strain DSM 11312 which was deposited on 02.12.96 (see WO 98/28409);

5 Fig. 7

this figure equals figure 2 of EP 0684313 and shows the amino acid and DNA sequence of a phytase ("M_thermophila") derived from strain ATCC 48102 (=ATCC 74340) of *Myceliophthora thermophila* which was re-deposited on 14.03.97;

10

Fig. 8

this figure shows the amino acid and DNA sequence of a phytase ("A_fumigatus") derived from strain ATCC 13073 of *Aspergillus fumigatus* (see EP 0897985);

15 Fig. 9

this figure shows the amino acid ("Conphys") and DNA sequence of an ascomycete consensus phytase (in the present context called "consphyA") (see EP 0897985);

Fig. 10

20

this figure shows the amino acid and DNA sequence of a phytase ("A_nidulans") derived from strain DSM 9743 of *Aspergillus nidulans* (see EP 0897985);

Fig. 11

25

this figure equals figure 8 of EP 0420358 and shows the amino acid and DNA sequence of a phytase ("A_ficum") derived from *Aspergillus ficum* strain NRRL-3135;

Fig. 12

30

this figure equals figure 1 of EP 0684313 and shows the amino acid and DNA sequence of a phytase ("A_terreus") derived from strain CBS 220.95 of *Aspergillus terreus*;

- Fig. 13 this figure shows the amino acid and DNA sequence of
a phytase ("T_thermo") derived from strain ATCC
20186 (=ATCC 74338) of *Talaromyces thermophilus*
which was redeposited on 14.03.97 (see EP 0897985);
- Fig. 14 this figure equals figure 2 of WO 97/35017 and shows
the amino acid and DNA sequence of a phytase
("T_lanuginosa") derived from strain CBS 586.94 of
Thermomyces lanuginosus; a plasmid comprising the
full length cDNA sequence was transformed into
E. coli DH5 α (pMWR46) strain B-21527 which was
deposited with NRRL on 23.02.96;
- Fig. 15 this figure shows the amino acid and DNA sequence of
a phytase ("C_foecundissimum") derived from strain
CBS 427.97 of *Cladorrhinum foecundissimum* which was
deposited on 23 January 1997; the expression plasmid
pYES 2.0 comprising the full length cDNA sequence
was transformed into E. coli strain DSM 12742 which
was deposited on 17 March 1999;
- Fig. 16 this figure shows an alignment of the phytase
C_foecundissimum with the model phytase
M_thermophila, using the program GAP gcg (Gap Weight
3.000; Length Weight 0.100); and
- Fig. 17 shows how the C_foecundissimum phytase can be pasted
onto the alignment of Fig. 1.

DETAILED DISCLOSURE OF THE INVENTION

Phytase

In the present context a phytase is an enzyme which catalyzes the hydrolysis of phytate (myo-inositol hexakisphosphate) to (1) myo-inositol and/or (2) mono-, di-, tri-, tetra- and/or penta-phosphates thereof and (3) inorganic phosphate. In the following, for short, the above compounds are sometimes referred to as IP6, I, IP1, IP2, IP3, IP4, IP5 and P, respectively. This means that by action of a phytase, IP6 is degraded into P + one or more of the components IP5, IP4, IP3, IP2, IP1 and I. Alternatively, myo-inositol carrying in total n phosphate groups attached to positions p, q, r,... is denoted Ins(p,q,r,...)Pn. For convenience Ins(1,2,3,4,5,6)P6 (phytic acid) is abbreviated PA.

According to the Enzyme nomenclature database ExPASy (a repository of information relative to the nomenclature of enzymes primarily based on the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB) describing each type of characterized enzyme for which an EC (Enzyme Commission) number has been provided), two different types of phytases are known: A so-called 3-phytase (myo-inositol hexaphosphate 3-phosphohydrolase, EC 3.1.3.8) and a so-called 6-phytase (myo-inositol hexaphosphate 6-phosphohydrolase, EC 3.1.3.26). The 3-phytase hydrolyses first the ester bond at the D-3-position, whereas the 6-phytase hydrolyzes first the ester bond at the D-6- or L-6-position.

The expression "phytase" or "polypeptide or enzyme exhibiting phytase activity" is intended to cover any enzyme capable of effecting the liberation of inorganic phosphate or phosphorous from various myo-inositol phosphates. Examples of

such myo-inositol phosphates (phytase substrates) are phytic acid and any salt thereof, e.g. sodium phytate or potassium phytate or mixed salts. Also any stereoisomer of the mono-, di-, tri-, tetra- or penta-phosphates of myo-inositol might serve as
5 a phytase substrate. A preferred phytase substrate is phytic acid and salts thereof.

In accordance with the above definition, the phytase activity can be determined using any assay in which one of these substrates is used. In the present context (unless otherwise
10 specified) the phytase activity is determined in the unit of FYT, one FYT being the amount of enzyme that liberates 1 μmol inorganic ortho-phosphate per min. under the following conditions: pH 5.5; temperature 37°C; substrate: sodium phytate ($\text{C}_6\text{H}_6\text{O}_{24}\text{P}_6\text{Na}_{12}$) in a concentration of 0.0050 mol/l. A suitable
15 phytase assay is described in the experimental part.

The present invention provides a genetically engineered phytase as described in the appending claims.

A genetically engineered phytase is a non-naturally occurring phytase which is different from a model phytase, e.g. a
20 wild-type phytase. Genetically engineered phytases include, but are not limited to, phytases prepared by site-directed mutagenesis, gene shuffling, random mutagenesis etc.

The invention also provides DNA constructs, vectors, host cells, and methods of producing these genetically engineered
25 phytases and phytase variants, as well as uses thereof.

A phytase variant is a polypeptide or enzyme or a fragment thereof which exhibits phytase activity and which is amended as compared to a model phytase.

Amended means altered by way of one or more amino acid or
30 peptide substitutions, deletions, insertions and/or additions - in each case by, or of, one or more amino acids. Such

substitutions, deletions, insertions, additions can be achieved by any method known in the art, e.g. gene shuffling, random mutagenesis, site-directed mutagenesis etc.

The model or parent phytase, from which the phytase variant is derived, can be any phytase, e.g. a wild-type phytase or a derivative, mutant or variant thereof, including allelic and species variants, as well as genetically engineered variants thereof, which e.g. can be prepared by site-directed mutagenesis, random mutagenesis, shuffling etc.

10 Included in the concept of model phytase is also any hybrid or chimeric phytase, i.e. a phytase which comprises a combination of partial amino acid sequences derived from at least two phytases.

The hybrid phytase may comprise a combination of partial
15 amino acid sequences deriving from at least two ascomycete phytases, at least two basidiomycete phytases or from at least one ascomycete and at least one basidiomycete phytase. These ascomycete and basidiomycete phytases from which a partial amino acid sequence derives may, e.g., be any of those specific
20 phytases referred to herein.

In the present context, a hybrid, shuffled, random mutagenised, site-directed mutagenised or otherwise genetically engineered phytase derived from ascomycete phytases only is also an ascomycete phytase; and a hybrid, shuffled, random
25 mutagenised, site-directed mutagenised or otherwise genetically engineered phytase derived from model basidiomycete phytases only is also a basidiomycete phytase. Any hybrid derived from at least one ascomycete phytase as well as at least one basidiomycete phytase is called a mixed ascomycete/basidiomycete
30 phytase and such phytase is also a model phytase in the present context.

Analogously, a hybrid, shuffled, random mutagenised, site-directed mutagenised or otherwise genetically engineered phytase derived from one or more *Aspergillus* phytases is also an *Aspergillus* derived phytase; and a hybrid, shuffled, random
5 mutagenised, site-directed mutagenised or otherwise genetically engineered phytase derived from any other taxonomic sub-grouping mentioned herein is also to be designated a phytase derived from this taxonomic sub-grouping.

Still further, in the present context, "derived from" is
10 intended to indicate a phytase produced or producible by a strain of the organism in question, but also a phytase encoded by a DNA sequence isolated from such strain and produced in a host organism transformed with said DNA sequence. Finally, the term is intended to indicate a phytase which is encoded by a DNA
15 sequence of synthetic and/or cDNA origin and which has the identifying characteristics of the phytase in question.

Preferably the model phytase is a phytase which can be aligned as described below to either of the thirteen phytases of Fig. 1 (which are particularly preferred model phytases).

20 Preferred wild-type model phytases (i.e. neither recombinant, or shuffled or otherwise genetically engineered phytases) have a degree of similarity or homology, preferably identity, to amino acid sequence no. 38-403 (*Peniophora* numbers) of either of these thirteen phytases of at least 40%, more
25 preferably at least 50%, still more preferably at least 60%, in particular at least 70%, especially at least 80%, and in a most preferred embodiment a degree of similarity of at least 90%.

Preferred recombinant or shuffled or otherwise genetically engineered model phytases have a degree of similarity or
30 homology, preferably identity, to partial sequence no. 38-49, 63-77, 274-291, 281-300 and 389-403 (*Peniophora* numbers) of

either of these thirteen phytases of at least 60%, more preferably at least 70%, still more preferably at least 80%, in particular at least 90%.

In a preferred embodiment the degree of similarity is based on a comparison with the complete amino acid sequence of either of the thirteen phytases.

The degree of similarity or homology, alternatively identity, can be determined using any alignment programme known in the art. A preferred alignment programme is GAP provided in the GCG version 8 program package (Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711) (see also Needleman, S.B. and Wunsch, C.D., (1970), Journal of Molecular Biology, 48, 443-453). Using GAP with the following settings for polypeptide sequence comparison: GAP weight of 3.000 and GAP lengthweight of 0.100.

Also preferred is a wild-type model phytase which comprises an amino acid sequence encoded by a DNA sequence which hybridizes to a DNA sequence encoding amino acid sequence 38-403 (Peniophora numbers) of any of the DNA sequences encoding the thirteen specific phytase sequences of Fig. 1.

A further preferred model phytase is a genetically engineered phytase, which comprises an amino acid sequence encoded by a DNA sequence which hybridizes to a DNA sequence encoding amino acid sequence 38-49, and to a DNA sequence encoding amino acid sequence 63-77, and to a DNA sequence encoding amino acid sequence 274-291, and to a DNA sequence encoding amino acid sequence 281-300, and to a DNA sequence encoding amino acid sequence 389-403 (Peniophora numbers) of any of the DNA sequences encoding the thirteen specific phytase sequences of Fig. 1.

In a preferred embodiment the hybridization is to the complete phytase encoding part of any of the thirteen phytases.

Suitable experimental conditions for determining whether a given DNA or RNA sequence "hybridizes" to a specified nucleotide or oligonucleotide probe involves presoaking of the filter containing the DNA fragments or RNA to examine for hybridization in 5 x SSC (Sodium chloride/Sodium citrate), (J. Sambrook, E.F. Fritsch, and T. Maniatis, 1989, Molecular Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, New York) for 10 min, and prehybridization of the filter in a solution of 5 x SSC, 5 x Denhardt's solution (Sambrook et al. 1989), 0.5 % SDS and 100 µg/ml of denatured sonicated salmon sperm DNA (Sambrook et al. 1989), followed by hybridization in the same solution containing a concentration of 10 ng/ml of a random-primed (Feinberg, A. P. and Vogelstein, B. (1983) Anal. Biochem. 132:6-13), 32P-dCTP-labeled (specific activity > 1 x 10⁹ cpm/µg) probe for 12 hours at approximately 45°C.

The filter is then washed twice for 30 minutes in 2 x SSC, 0.5 % SDS at at least 55°C (low stringency), at at least 60°C (medium stringency), at at least 65°C (medium/high stringency), at at least 70°C (high stringency), or at at least 75°C (very high stringency).

Molecules to which the oligonucleotide probe hybridizes under these conditions are detected using an x-ray film.

It should be noted that a certain specific phytase variant need not actually have been prepared from a specific model phytase, for this model phytase to qualify as a "model phytase" in the present context. It is sufficient that the variant exhibits at least one of the herein indicated amendments when it is afterwards compared with the model phytase.

The alignment of Fig. 1 is made using the program PileUp (Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711), with a GapWeight of 3.000 and a
5 GapLengthWeight of 0.100. When aligning a new model phytase or a new phytase variant all thirteen sequences can be included together with the new phytase (variant) in a multiple alignment, or, alternatively, at least one of the thirteen sequences of Fig. 1 is included together with the new phytase (variant) in an
10 alignment.

A preferred procedure for aligning according to Fig. 1 a new model phytase (or a phytase variant) is as follows: The new model phytase is aligned with that specific sequence of the thirteen sequences of Fig. 1 to which the new model phytase has
15 the highest degree of homology. For calculating the degree of homology, and for making the "alignment according to Fig. 1" of the two sequences, the program GAP referred to below is preferably used. Having aligned the two sequences, the new model phytase (or phytase variant) is added (pasted) to the alignment
20 at Fig. 1 using the result of the first alignment (placing identical and homologous amino acid residues above each other as prescribed by the alignment), following which corresponding positions are now easily identifiable.

Example 7 shows an example of how to add a new model
25 phytase to the alignment of Fig. 1 and deduce corresponding phytase variants thereof.

Other model phytases can be aligned and variants deduced in analogy with Example 7. This is so in particular for the following model phytases: The phytase of *Aspergillus niger* var.
30 awamori (US patent no. 5,830,733); the *Bacillus* phytase of WO 98/06858; the soy bean phytase of WO 98/20139; the maize

phytase of WO 98/05785; the *Aspergillus* phytase of WO 97/38096; the phytases of *Monascus anka* of WO 98/13480; the phytase from *Schwanniomyces occidentalis* of EP 0699762 etc.

When comparing a model phytase and a proposed phytase
5 variant using the alignment as described herein, corresponding amino acid positions can be identified, viz. a model position of the model phytase and a variant position of the variant - the corresponding model position and variant position are simply placed one above the other in the alignment. An amendment is
10 said to have occurred in a given position if the model amino acid of the model position and the variant amino acid of the variant position are different. Preferred amendments of these positions manifest themselves as amino acid substitutions, deletions or additions.

15 Amended in at least one position means amended in one or more positions, i.e. in one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve etc. up to all N positions listed. This definition includes any possible sub-combinations thereof, e.g. any set of two substitutions, any set of three,
20 any set of four, etc. - to any set of (N-1) positions.

In the present context all sequences, whatever the model phytase, and including the thirteen sequences of Fig. 1, are numbered using the numbering corresponding to the phytase *P_lycii*. These "Peniophora numbers" are indicated at Fig. 1,
25 together with the "alignment numbers." The numbering of *P_lycii* starts at M1 and ends at E439.

As explained above, the alignment reveals which positions in various phytase sequences other than *P_lycii* are equivalent or corresponding to the given *P. lycii* position.

30 A substitution of amino acids is indicated herein as for instance "3S," which indicates, that at position 3 amino acid S

should be substituted for the "original" or model position 3 amino acid, whichever it is. Thus, the substitution should result in an S in the corresponding variant position. Considering now the alignment at Fig. 1, a substitution like
 5 e.g. "3S" is to be interpreted as follows, for the respective phytases shown (the amino acid first indicated is the "original" or model amino acid in "Peniophora position" 3):

	P_involtus_A1:	F3S (number 3 F substituted by S)
	P_involtus_A2:	L3S
10	T_pubescens:	M1S
	A_pediades:	M1S
	P_lycii:	redundant (already an S)
	A_fumigatus:	T5S
	consphyA:	V5S
15	A_nidulans:	T5S
	A_ficum_NRRL3135:	A5S
	A_terreus:	A5S
	T_thermo:	L5S
	T_lanuginosa:	V11S
20	M_thermophila:	G5S

However, in what follows the above specific substitutions will be designated as follows (always using the Peniophora numbering):

25	P_involtus_A1:	F3S
	P_involtus_A2:	L3S
	T_pubescens:	M3S
	A_pediades:	M3S
	P_lycii:	redundant (already an S)
30	A_fumigatus:	T3S
	consphyA:	V3S

16

	A_nidulans:	T3S
	A_ficuum_NRRL3135:	A3S
	A_terreus:	A3S
	T_thermo:	L3S
5	T_lanuginosa:	V3S
	M_thermophila:	G3S

Still further, denotations like e.g. "3S,F,G" means that the amino acid in position 3 (Peniophora numbers) of the model
 10 phytase in question is substituted with either of S, F or G, i.e. e.g. the designation "3S,F,G" is considered fully equivalent to the designation "3S, 3F, 3G".

A denotation like ()3S means that amino acid S is added to the sequence in question (at a gap in the actual sequence), in a
 15 position corresponding to Peniophora number 3 - and vice versa for deletions (S3()).

In case of regions in which the Peniophora phytase sequence has larger deletions than some of the other phytases in Fig. 1, for instance in the region between position 201 and 202
 20 (Peniophora numbers), intermediate positions (amino acid residues in other sequences) are numbered by adding a,b,c,d, etc, in lower-case letters, to the last Peniophora position number, e.g. for the phytase M_thermophila: E201; G201a; P201b; Y201c; S201d; T201e; I201f; G202; D203 etc.

15 In one of the priority applications of the present application there are two minor position numbering errors: According to the above definitions, the positions referred to in first priority application as 204 and 205 (Peniophora
 ers) are wrongly designated; they should have been numbered
 nd 204, respectively. Therefore, 204 has been substituted
 and 205 by 204 throughout the present application.

A preferred phytase variant of the invention comprises an amino acid sequence which comprises, preferably contains, one or more of the following amino acid substitutions:

24C; 27P; 31Y; 33C; 39H,S,Q; 40L,N; 42S,G;
 5 43A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y; 44N; 45D,S; 47Y,F;
 49P; 51E,A,R; 56P; 58D,K,A; 59G; 61R; 62V,I; 69Q; 75W,F; 78D,S;
 79G; 80K,A; 81A,G,Q,E; 82T; 83A,I,K,R,Q; 84I,Y,Q,V; 88I; 90R,A;
 102Y; 115N; 116S; 118V,L; 119E; 120L; 122A; 123N,Q,T; 125M,S;
 126H,S,V; 127Q,E,N; 128A,S,T; 132F,I,L; 143N; 148V,I; 151A,S;
 10 152G; 153D,Y; 154D,Q,S,G; 157V; 158D,A; 159T; 160A,S; 161T,N;
 162N; 163W; 170FH; 170gA; 171N; 172P; 173Q,S; 184Q,S,P; 185S;
 186A,E,P; 187A; 187aS; 190A,P; 193S; 194S,T; 195T,V,L; 198A,N,V;
 200G,V; 201D,E; a deletion of at least one of 201a, 201b, 201c,
 201d, 201e, 201f, preferably all; 201eT; 202S,A; 203R,K,S;
 15 203aV,T; 204Q,E,S,A,V; 205E; 211L,V; 215A,P; 220L,N; 223H,D;
 228N; 232T; 233E; 235Y,L,T; 236Y,N; 237F; 238L,M; 242P,S; 244D;
 246V; 251eE,Q; 253P; 256D; 260A,H; 264R,I; 265A,Q; 267D;
 270Y,A,L,G; 271D,N; 273D,K; 275F,Y; 278T,H; 280A,P; 283P;
 287A,T; 288L,I,F; 292F,Y; 293A,V; 302R,H; 304P,A; 332F; 336S;
 20 337T,G,Q,S; 338I; 339V,I; 340P,A; 343A,S,F,I,L; 348Y; 349P;
 352K; 360R; 362P; 364W,F; 365V,L,A,S; 366D,S,V; 367A,K; 368K;
 369I,L; 370V; 373A,S; 374S,A; 375H; 376M; 383kQ,E; 387P; 393V;
 396R; 404A,G; 409R; 411K,T; 412R; 417E,R; 421F,Y; 431E.

In a preferred embodiment this is with the proviso that
 25 the model phytase does not already comprise the above suggested
 amino acid substitution or addition or deletion at the position
 indicated. Or, with the proviso that, for each position, the
 model amino acid is not already the variant amino acid hereby
 proposed. But these provisos can be said to be in fact already
 30 inherent in the above wording, because of the expression
 "amended."

The various preferred phytase variants of claims 16-34 comprises, preferably contains or have, amino acid sequences which comprise or contain one or more of the amino acid substitutions, additions, or deletions listed in the respective 5 claims.

In a preferred embodiment the various phytase variants comprise 1, 2, 3, 4, 5, 6, 7, 8, 9 or even 10 of these substitutions; or a number of substitutions of 10-15, 15-20, 20-30 or even 30-50; eventually up to 60, 70, 80 or 90 10 substitutions.

In another preferred embodiment, the amino acid sequence of the various phytase variants comprise one or more substitutions of the substitution sub-groupings listed hereinbelow; or combinations of substitutions classified in two 15 or more sub-groupings.

Generally, instead of "comprise," "contain" or "have," the amino acid sequences of preferred variants "consist essentially of" or "consist of" the specific model phytases of fig. 1, as modified by one or more of the substitutions described herein.

20 In the present context a basidiomycete means a microorganism of the phylum Basidiomycota. This phylum of Basidiomycota is comprised in the fungal kingdom together with e.g. the phylum Ascomycota ("ascomycetes").

Taxonomical questions can be clarified by consulting the 25 references listed below or by consulting a fungal taxonomy database (NIH Data Base (Entrez)) which is available via the Internet on World Wide Web at the following address: <http://www3.ncbi.nlm.nih.gov/Taxonomy/tax.html>.

For a definition of basidiomycetes, reference is made to 30 either Jülich, 1981, Higher Taxa of Basidiomycetes; Ainsworth & Bisby's (eds.) Dictionary of the Fungi, 1995, Hawksworth, D.L.,

P.M. Kirk, B.C. Sutton & D.N. Pegler; or Hansen & Knudsen (Eds.), Nordic Macromycetes, vol. 2 (1992) and 3 (1997). A preferred reference is Hansen & Knudsen.

For a definition of ascomycetes, reference is made to either of Ainsworth & Brisby cited above or Systema Ascomycetum by Eriksson, O.E. & D. L. Hawksworth, Vol. 16, 1998. A preferred reference is Eriksson et al.

Generally, a microorganism which is classified as a basidiomycete/ascomycete in either of the references listed above, including the database, is a basidiomycete/ascomycete in the present context.

Some *Aspergillus* strains are difficult to classify because they are anamorphous, and therefore they might be classified in Fungi Imperfecti. However, once the teleomorphous counterpart is found, it is re-classified taxonomically. For instance, the teleomorph of *A. nidulans* is *Emericella nidulans* (of the family Trichocomaceae, the order Eurotiales, the class Plectomycetes of the phylum Ascomycota). These subgroupings of Ascomycota are preferred, together with the family Lasiosphaeriaceae, the order Sordariales, the class Pyrenomycetes of the phylum Ascomycota.

The wording "ascomycetes" and analogues as used herein includes any strains of *Aspergillus*, *Thermomyces*, *Myceliophthora*, *Talaromyces* which are anamorphous and thus would be classified in Fungi Imperfecti.

Preferred basidiomycete phytases are those listed in WO 98/28409, in the very beginning of the section headed "Detailed description of the invention".

DNA sequences encoding the thirteen specifically listed model phytases and other model phytases can be prepared according to the teachings of each of the documents listed under the brief description of the drawings.

A DNA sequence encoding a model phytase may be isolated from any cell or microorganism producing the phytase in question, using various methods well known in the art. First, a genomic DNA and/or cDNA library should be constructed using
5 chromosomal DNA or messenger RNA from the organism that produces the phytase. Then, if the amino acid sequence of the phytase is known, homologous, labelled oligonucleotide probes may be synthesized and used to identify phytase-encoding clones from a genomic library prepared from the organism in question.
10 Alternatively, a labelled oligonucleotide probe containing sequences homologous to a known phytase gene could be used as a probe to identify phytase-encoding clones, using hybridization and washing conditions of lower stringency.

Yet another method for identifying phytaseencoding clones
15 would involve inserting fragments of genomic DNA into an expression vector, such as a plasmid, transforming phytase-negative bacteria with the resulting genomic DNA library, and then plating the transformed bacteria onto agar containing a substrate for phytase thereby allowing clones expressing the
20 phytase to be identified.

Alternatively, the DNA sequence encoding the enzyme may be prepared synthetically by established standard methods, e.g. the phosphoroamidite method described by S.L. Beaucage and M.H. Caruthers (1981) or the method described by Matthes et al.
25 (1984). In the phosphoroamidite method, oligonucleotides are synthesized, e.g. in an automatic DNA synthesizer, purified, annealed, ligated and cloned in appropriate vectors.

Finally, the DNA sequence may be of mixed genomic and synthetic origin, mixed synthetic and cDNA origin or mixed genomic
30 and cDNA origin, prepared by ligating fragments of synthetic, genomic or cDNA origin (as appropriate, the fragments

corresponding to various parts of the entire DNA sequence), in accordance with standard techniques. The DNA sequence may also be prepared by polymerase chain reaction (PCR) using specific primers, for instance as described in US 4,683,202 or R.K. Saiki et al. (1988).

DNA encoding the phytase variants of the present invention can be prepared by methods known in the art, such as Site-directed Mutagenesis. Once a DNA sequence encoding a model phytase of interest has been isolated, and desirable sites for mutation identified, mutations may be introduced using synthetic oligonucleotides. These oligonucleotides contain nucleotide sequences flanking the desired mutation sites; mutant nucleotides are inserted during oligonucleotide synthesis. In a specific method, a single-stranded gap of DNA, bridging the phytase-encoding sequence, is created in a vector carrying the phytase-encoding gene. Then the synthetic nucleotide, bearing the desired mutation, is annealed to a homologous portion of the single-stranded DNA. The remaining gap is then filled in with DNA polymerase I (Klenow fragment) and the construct is ligated using T4 ligase. A specific example of this method is described in Morinaga et al. (1984). US 4,760,025 discloses the introduction of oligonucleotides encoding multiple mutations by performing minor alterations of the cassette. However, an even greater variety of mutations can be introduced at any one time by the Morinaga method because a multitude of oligonucleotides, of various lengths, can be introduced.

Another method of introducing mutations into DNA sequences encoding a desired model phytase is described in Nelson and Long (1989). It involves a 3-step generation of a PCR fragment containing the desired mutation introduced by using a chemically synthesized DNA strand as one of the primers in the PCR

reactions. From the PCR-generated fragment, a DNA fragment carrying the mutation may be isolated by cleavage with restriction endonucleases and reinserted into an expression plasmid.

5 Yet another method of mutating DNA sequences encoding a model phytase is Random Mutagenesis. Random mutagenesis is suitably performed either as localised or region-specific random mutagenesis in at least three parts of the gene translating to the amino acid sequence shown in question, or within the whole
10 gene.

The random mutagenesis of a DNA sequence encoding a model phytase may be conveniently performed by use of any method known in the art.

In relation to the above, further aspects of the present
15 invention relates to a method for generating a variant of a model phytase, wherein the variant preferably exhibits amended characteristics as described below, the method comprising:

(a) subjecting a DNA sequence encoding the model phytase to Site-directed Mutagenesis, or the Nelson and Long PCR
20 mutagenesis method or to Random Mutagenesis,

(b) expressing the mutated DNA sequence obtained in step (a) in a host cell, and

(c) screening for host cells expressing a phytase variant which has an altered property relative to the model
25 phytase.

When using Random Mutagenesis, step (a) of the above method of the invention is preferably performed using doped primers.

For instance, the random mutagenesis may be performed by
30 use of a suitable physical or chemical mutagenizing agent, by use of a suitable oligonucleotide, or by subjecting the DNA

sequence to PCR generated mutagenesis. Furthermore, the random mutagenesis may be performed by use of any combination of these mutagenizing agents. The mutagenizing agent may, e.g., be one which induces transitions, transversions, inversions, scrambling, deletions, and/or insertions.

Examples of a physical or chemical mutagenizing agent suitable for the present purpose include ultraviolet (UV) irradiation, hydroxylamine, N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), O-methyl hydroxylamine, nitrous acid, ethyl methane sulphonate (EMS), sodium bisulphite, formic acid, and nucleotide analogues. When such agents are used, the mutagenesis is typically performed by incubating the DNA sequence encoding the parent enzyme to be mutagenized in the presence of the mutagenizing agent of choice under suitable conditions for the mutagenesis to take place, and selecting for mutated DNA having the desired properties.

When the mutagenesis is performed by the use of an oligonucleotide, the oligonucleotide may be doped or spiked with the three non-parent nucleotides during the synthesis of the oligonucleotide at the positions which are to be changed. The doping or spiking may be done so that codons for unwanted amino acids are avoided. The doped or spiked oligonucleotide can be incorporated into the DNA encoding the phytase enzyme by any published technique, using e.g. PCR, LCR or any DNA polymerase and ligase as deemed appropriate.

Preferably, the doping is carried out using "constant random doping", in which the percentage of wild-type and mutation in each position is predefined. Furthermore, the doping may be directed toward a preference for the introduction of certain nucleotides, and thereby a preference for the introduction of one or more specific amino acid residues. The

doping may be made, e.g., so as to allow for the introduction of 90% wild type and 10% mutations in each position. An additional consideration in the choice of a doping scheme is based on genetic as well as protein-structural constraints. The
5 doping scheme may be made by using the DOPE program which, inter alia, ensures that introduction of stop codons is avoided.

When PCR-generated mutagenesis is used, either a chemically treated or non-treated gene encoding a model phytase is subjected to PCR under conditions that increase the mis-
10 incorporation of nucleotides (Deshler 1992; Leung et al., Technique, Vol.1, 1989, pp. 11-15).

A mutator strain of *E. coli* (Fowler et al., Molec. Gen. Genet., 133, 1974, pp. 179-191), *S. cerevisiae* or any other microbial organism may be used for the random mutagenesis of the
15 DNA encoding the model phytase by, e.g., transforming a plasmid containing the parent glycosylase into the mutator strain, growing the mutator strain with the plasmid and isolating the mutated plasmid from the mutator strain. The mutated plasmid may be subsequently transformed into the expression organism.

20 The DNA sequence to be mutagenized may be conveniently present in a genomic or cDNA library prepared from an organism expressing the model phytase. Alternatively, the DNA sequence may be present on a suitable vector such as a plasmid or a bacteriophage, which as such may be incubated with or otherwise
25 exposed to the mutagenising agent. The DNA to be mutagenized may also be present in a host cell either by being integrated in the genome of said cell or by being present on a vector harboured in the cell. Finally, the DNA to be mutagenized may be in isolated form. It will be understood that the DNA
30 sequence to be subjected to random mutagenesis is preferably a cDNA or a genomic DNA sequence.

In some cases it may be convenient to amplify the mutated DNA sequence prior to performing the expression step b) or the screening step c). Such amplification may be performed in accordance with methods known in the art, the presently preferred method being PCR-generated amplification using oligonucleotide primers prepared on the basis of the DNA or amino acid sequence of the parent enzyme.

Subsequent to the incubation with or exposure to the mutagenising agent, the mutated DNA is expressed by culturing a suitable host cell carrying the DNA sequence under conditions allowing expression to take place. The host cell used for this purpose may be one which has been transformed with the mutated DNA sequence, optionally present on a vector, or one which was carried the DNA sequence encoding the parent enzyme during the mutagenesis treatment. Examples of suitable host cells are the following: gram positive bacteria such as *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus brevis*, *Bacillus stearothermophilus*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus circulans*, *Bacillus lautus*, *Bacillus megaterium*, *Bacillus thuringiensis*, *Streptomyces lividans* or *Streptomyces murinus*; and gram-negative bacteria such as *E. coli*.

The mutated DNA sequence may further comprise a DNA sequence encoding functions permitting expression of the mutated DNA sequence.

The random mutagenesis may be advantageously localised to a part of the model phytase in question using Localized random mutagenesis. This may, e.g., be advantageous when certain regions of the enzyme have been identified to be of particular importance for a given property of the enzyme, and when modified are expected to result in a variant having improved properties.

Such regions may normally be identified when the tertiary structure of the parent enzyme has been elucidated and related to the function of the enzyme.

The localized, or region-specific, random mutagenesis is conveniently performed by use of PCR generated mutagenesis techniques as described above or any other suitable technique known in the art. Alternatively, the DNA sequence encoding the part of the DNA sequence to be modified may be isolated, e.g., by insertion into a suitable vector, and said part may be subsequently subjected to mutagenesis by use of any of the mutagenesis methods discussed above.

For region-specific random mutagenesis with a view to amending e.g. the specific activity of a model phytase, codon positions corresponding to the following amino acid residues from the amino acid sequences set forth in Fig. 1 may appropriately be targeted:

Residues: 41-47, 68-80, 83-84, 115-118, 120-126, 128, 149-163, 184-185, 191-193, 198-201e, 202-203, 205, 235-236, 238-239, 242-243, 270-279, 285, 288, 332-343, 364-367, 369-375, 394.

Regions: 41-47, 68-80, 120-128, 149-163, 270-279, 332-343, 364-375.

The random mutagenesis may be carried out by the following steps:

1. Select regions of interest for modification in the parent enzyme
2. Decide on mutation sites and non-mutated sites in the selected region
3. Decide on which kind of mutations should be carried out, e.g. with respect to the desired stability and/or performance of the variant to be constructed
4. Select structurally reasonable mutations

5. Adjust the residues selected by step 3 with regard to step 4.

6. Analyse by use of a suitable dope algorithm the nucleotide distribution.

5 7. If necessary, adjust the wanted residues to genetic code realism, e.g. taking into account constraints resulting from the genetic code, e.g. in order to avoid introduction of stop codons; the skilled person will be aware that some codon combinations cannot be used in practice and will need to be
10 adapted

8. Make primers

9. Perform random mutagenesis by use of the primers

10. Select resulting phytase variants by screening for the desired improved properties.

15 Suitable dope algorithms for use in step 6 are well known in the art. One such algorithm is described by Tomandl, D. et al., 1997, Journal of Computer-Aided Molecular Design 11:29-38. Another algorithm is DOPE (Jensen, LJ, Andersen, KV, Svendsen, A, and Kretzschmar, T (1998) Nucleic Acids Research 26:697-702).

20 A DNA sequence encoding a model phytase or a phytase variant of the invention can be expressed using an expression vector, a recombinant expression vector, which typically includes control sequences encoding a promoter, operator, ribosome binding site, translation initiation signal, and,
25 optionally, a repressor gene or various activator genes.

The recombinant expression vector may be any vector which may conveniently be subjected to recombinant DNA procedures, and the choice of vector will often depend on the host cell into which it is to be introduced. Thus, the vector may be an
30 autonomously replicating vector, e.g. a plasmid, a bacteriophage or an extra-chromosomal element. Alternatively, the vector may

be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated.

In the vector, the DNA sequence should be operably
5 connected to a suitable promoter sequence. The promoter may be any DNA sequence which shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell. An example of a suitable promoter for directing the transcription
10 of the DNA sequence encoding a phytase variant of the invention, especially in a bacterial host, is the promoter of the lac operon of E.coli. For transcription in a fungal host, examples of useful promoters are those derived from the gene encoding A. oryzae TAKA amylase.

15 The expression vector of the invention may also comprise a suitable transcription terminator and, in eukaryotes, polyadenylation sequences operably connected to the DNA sequence encoding the phytase variant of the invention. Termination and polyadenylation sequences may suitably be derived from the same
20 sources as the promoter.

The vector may further comprise a DNA sequence enabling the vector to replicate in the host cell in question. Examples of such sequences are the origins of replication of plasmids pUC19, pACYC177, pUB110, pE194, pAMB1 and pIJ702.

25 The vector may also comprise a selectable marker, e.g. a gene the product of which complements a defect in the host cell, such as the dal genes from B. subtilis or B. licheniformis, or one which confers antibiotic resistance such as ampicillin resistance. Furthermore, the vector may comprise Aspergillus
30 selection markers such as amdS, argB, niaD and sC, or the

selection may be accomplished by co-transformation, e.g. as described in WO 91/17243.

The procedures used to ligate the DNA construct of the invention encoding a phytase variant, the promoter, terminator
5 and other elements, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook et al. (1989)).

The cell of the invention, either comprising a DNA
10 construct or an expression vector of the invention as defined above, is advantageously used as a host cell in the recombinant production of a phytase variant of the invention. The cell may be transformed with the DNA construct of the invention encoding the variant, conveniently by integrating the DNA construct (in
15 one or more copies) in the host chromosome. This integration is generally considered to be an advantage as the DNA sequence is more likely to be stably maintained in the cell. Integration of the DNA constructs into the host chromosome may be performed according to conventional methods, e.g. by homologous or
20 heterologous recombination. Alternatively, the cell may be transformed with an expression vector as described above in connection with the different types of host cells.

An isolated DNA molecule or, alternatively, a "cloned DNA sequence" "a DNA construct," "a DNA segment" or "an isolated DNA
25 sequence" refers to a DNA molecule or sequence which can be cloned in accordance with standard cloning procedures used in genetic engineering to relocate the DNA segment from its natural location to a different site where it will be replicated. The term refers generally to a nucleic acid sequence which is
30 essentially free of other nucleic acid sequences, e.g., at least about 20% pure, preferably at least about 40% pure, more

preferably about 60% pure, even more preferably about 80% pure, most preferably about 90% pure, and even most preferably about 95% pure, as determined by agarose gel electrophoresis. The cloning procedures may involve excision and isolation of a
5 desired nucleic acid fragment comprising the nucleic acid sequence encoding the polypeptide, insertion of the fragment into a vector molecule, and incorporation of the recombinant vector into a host cell where multiple copies or clones of the nucleic acid sequence will be replicated. The nucleic acid
10 sequence may be of genomic, cDNA, RNA, semisynthetic, synthetic origin, or any combinations thereof.

The term "vector" is intended to include such terms/objects as "nucleic acid constructs," "DNA constructs," expression vectors" or "recombinant vectors."

15 The nucleic acid construct comprises a nucleic acid sequence of the present invention operably linked to one or more control sequences capable of directing the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences.

20 "Nucleic acid construct" is defined herein as a nucleic acid molecule, either single or double-stranded, which is isolated from a naturally occurring gene or which has been modified to contain segments of nucleic acid which are combined and juxtaposed in a manner which would not otherwise exist in
25 nature.

The term nucleic acid construct may be synonymous with the term expression cassette when the nucleic acid construct contains all the control sequences required for expression of a coding sequence of the present invention.

30 The term "coding sequence" as defined herein primarily comprises a sequence which is transcribed into mRNA and

translated into a polypeptide of the present invention when placed under the control of the above mentioned control sequences. The boundaries of the coding sequence are generally determined by a translation start codon ATG at the 5'-terminus 5 and a translation stop codon at the 3'-terminus. A coding sequence can include, but is not limited to, DNA, cDNA, and recombinant nucleic acid sequences.

The term "control sequences" is defined herein to include all components which are necessary or advantageous for 10 expression of the coding sequence of the nucleic acid sequence. Each control sequence may be native or foreign to the nucleic acid sequence encoding the polypeptide. Such control sequences include, but are not limited to, a leader, a polyadenylation sequence, a propeptide sequence, a promoter, a signal sequence, 15 and a transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences 20 with the coding region of the nucleic acid sequence encoding a polypeptide.

A "host cell" or "recombinant host cell" encompasses any progeny of a parent cell which is not identical to the parent cell due to mutations that occur during replication.

25 The cell is preferably transformed with a vector comprising a nucleic acid sequence of the invention followed by integration of the vector into the host chromosome.

"Transformation" means introducing a vector comprising a nucleic acid sequence of the present invention into a host cell 30 so that the vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector. Integration is

generally considered to be an advantage as the nucleic acid sequence is more likely to be stably maintained in the cell. Integration of the vector into the host chromosome may occur by homologous or non-homologous recombination as described above.

5 The host cell may be a unicellular microorganism, e.g., a prokaryote, or a non-unicellular microorganism, e.g., a eukaryote. Examples of a eukaryote cell is a mammalian cell, an insect cell, a plant cell or a fungal cell. Useful mammalian cells include Chinese hamster ovary (CHO) cells, HeLa cells,
10 baby hamster kidney (BHK) cells, COS cells, or any number of other immortalized cell lines available, e.g., from the American Type Culture Collection.

In a preferred embodiment, the host cell is a fungal cell.

Fungal cells may be transformed by a process involving
15 protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known per se.

The present invention also relates to a transgenic plant, plant part, such as a plant seed, or plant cell, which has been transformed with a DNA sequence encoding the phytase of the
20 invention so as to express or produce this enzyme. Also compositions and uses of such plant or plant part are within the scope of the invention, especially its use as feed and food or additives therefore, along the lines of the present use and food/feed claims.

25 The transgenic plant can be dicotyledonous or monocotyledonous, for short a dicot or a monocot. Of primary interest are such plants which are potential food or feed components and which comprise phytic acid. A normal phytic acid level of feed components is 0.1-100 g/kg, or more usually 0.5-50
30 g/kg, most usually 0.5-20 g/kg. Examples of monocot plants are grasses, such as meadow grass (blue grass, Poa), forage grass

such as festuca, lolium, temperate grass, such as Agrostis, and cereals, e.g. wheat, oats, rye, barley, rice, sorghum and maize (corn).

Examples of dicot plants are legumes, such as lupins, pea,
5 bean and soybean, and cruciferous (family Brassicaceae), such as cauliflower, oil seed rape and the closely related model organism Arabidopsis thaliana.

Such transgenic plant etc. is capable of degrading its own phytic acid, and accordingly the need for adding such enzymes to
10 food or feed comprising such plants is alleviated. Preferably, the plant or plant part, e.g. the seeds, are ground or milled, and possibly also soaked before being added to the food or feed or before the use, e.g. intake, thereof, with a view to adapting the speed of the enzymatic degradation to the actual use.

15 If desired, the plant produced enzyme can also be recovered from the plant. In certain cases the recovery from the plant is to be preferred with a view to securing a heat stable formulation in a potential subsequent pelleting process.

Examples of plant parts are stem, callus, leaves, root,
20 fruits, seeds, tubers etc. But also any plant tissue is included in this definition.

Any plant cell, whatever the tissue origin, is included in the definition of plant cells above.

Also included within the scope of the invention are the
25 progeny of such plants, plant parts and plant cells.

The skilled man will know how to construct a DNA expression construct for insertion into the plant in question, paying regard i.a. to whether the enzyme should be excreted in a tissue specific way. Of relevance for this evaluation is the
30 stability (pH-stability, degradability by endogenous proteases etc.) of the phytase in the expression compartments of the

plant. He will also be able to select appropriate regulatory sequences such as promoter and terminator sequences, and signal or transit sequences if required (Tague et al, Plant, Phys., 86, 506, 1988).

5 The plant, plant part etc. can be transformed with this DNA construct using any known method. An example of such method is the transformation by a viral or bacterial vector such as bacterial species of the genus *Agrobacterium* genetically engineered to comprise the gene encoding the phytase of the
10 invention. Also methods of directly introducing the phytase DNA into the plant cell or plant tissue are known in the art, e.g. micro injection and electroporation (Gasser et al, Science, 244, 1293; Potrykus, Bio/Techn. 8, 535, 1990; Shimamoto et al, Nature, 338, 274, 1989).

15 Following the transformation, the transformants are screened using any method known to the skilled man, following which they are regenerated into whole plants.

These plants etc. as well as their progeny then carry the phytase encoding DNA as a part of their genetic equipment.

20 In general, reference is made to WO 9114782A and WO 9114772A.

Agrobacterium tumefaciens mediated gene transfer is the method of choice for generating transgenic dicots (for review Hooykas & Schilperoort, 1992. Plant Mol. Biol. 19: 15-38),
25 however it can also be used for transforming monocots. Due to host range limitations it is generally not possible to transform monocots with the help of *A. tumefaciens*. Here, other methods have to be employed. The method of choice for generating transgenic monocots is particle bombardment (microscopic gold or
30 tungsten particles coated with the transforming DNA) of embryonic calli or developing embryos (Christou, 1992. Plant J.

2: 275-281; Shimamoto, 1994. *Curr. Opin. Biotechnol.* 5: 158-162; Vasil et al., 1992. *Bio/Technology* 10: 667-674).

Also other systems for the delivery of free DNA into these plants, including viral vectors (Joshi & Joshi, 1991. *FEBS Lett.* 5 281: 1-8), protoplast transformation via polyethylene glycol or electroporation (for review see Potyrkus, 1991. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42: 205-225), microinjection of DNA into mesophyll protoplasts (Crossway et al., 1986. *Mol. Gen. Genet.* 202: 79-85), and macroinjection of DNA into young floral 10 tillers of cereal plants (de la Pena et al., 1987. *Nature* 325: 274-276) are preferred methods.

In general, the cDNA or gene encoding the phytase variant of the invention is placed in an expression cassette (e.g. Pietrzak et al., 1986. *Nucleic Acids Res.* 14: 5857-5868) 15 consisting of a suitable promoter active in the target plant and a suitable terminator (termination of transcription). This cassette (of course including a suitable selection marker, see below) will be transformed into the plant as such in case of monocots via particle bombardment. In case of dicots the 20 expression cassette is placed first into a suitable vector providing the T-DNA borders and a suitable selection marker which in turn are transformed into *Agrobacterium tumefaciens*. Dicots will be transformed via the *Agrobacterium* harbouring the expression cassette and selection marker flanked by T-DNA 25 following standard protocols (e.g. Akama et al., 1992. *Plant Cell Reports* 12: 7-11). The transfer of T-DNA from *Agrobacterium* to the Plant cell has been recently reviewed (Zupan & Zambryski, 1995. *Plant Physiol.* 107: 1041-1047). Vectors for plant transformation via *Agrobacterium* are commercially available or 30 can be obtained from many labs that construct such vectors (e.g. Deblaere et al., 1985. *Nucleic Acids Res.* 13: 4777-4788; for

review see Klee et al., 1987. Annu. Rev. Plant Physiol. 38: 467-486).

Available plant promoters: Depending on the process under manipulation, organ- and/or cell-specific expression as well as appropriate developmental and environmental control may be required. For instance, it is desirable to express a phytase cDNA in maize endosperm etc. The most commonly used promoter has been the constitutive 35S-CaMV promoter Franck et al., 1980. Cell 21: 285-294). Expression will be more or less equal throughout the whole plant. This promoter has been used successfully to engineer herbicide- and pathogen-resistant plants (for review see Stitt & Sonnewald, 1995. Annu. Rev. Plant Physiol. Plant Mol. Biol. 46: 341-368). Organ-specific promoters have been reported for storage sink tissues such as seeds, potato tubers, and fruits (Edwards & Coruzzi, 1990. Annu. Rev. Genet. 24: 275-303), and for metabolic sink tissues such as meristems (Ito et al., 1994. Plant Mol. Biol. 24: 863-878).

The medium used to culture the transformed host cells may be any conventional medium suitable for growing the host cells in question. The expressed phytase may conveniently be secreted into the culture medium and may be recovered therefrom by well-known procedures including separating the cells from the medium by centrifugation or filtration, precipitating proteinaceous components of the medium by means of a salt such as ammonium sulphate, followed by chromatographic procedures such as ion exchange chromatography, affinity chromatography, or the like.

Preferred host cells are a strain of *Fusarium*, *Hansenula*, *Trichoderma* or *Aspergillus*, in particular a strain of *Fusarium graminearum*, *Fusarium venenatum*, *Fusarium cerealis*, *Fusarium* sp. having the identifying characteristic of *Fusarium* ATCC 20334, as further described in PCT/US/95/07743, *Hansenula polymorpha*,

Trichoderma harzianum or Trichoderma reesei, Aspergillus niger or Aspergillus oryzae.

References for expression in Hansenula polymorpha: Gellissen, G., Piontek, M., Dahlems, U., Jenzelewski, V.,
5 Gavagan, J.E., DiCosimo, R., Anton, D.I. & Janowicz, Z.A. (1996) Recombinant Hansenula polymorpha as a biocatalyst: coexpression of the spinach glycolate oxidase (GO) and the S. cerevisiae catalase T (CTT1) gene. Appl. Microbiol. Biotechnol. 46, 46-54.

Some more specific uses of the phytase variants according
10 to the invention appear from PCT/DK97/00568, the last pages of the detailed description of the invention section.

In a preferred embodiment, the phytase variant of the invention is essentially free of other non-phytase polypeptides, e.g., at least about 20% pure, preferably at least about 40%
15 pure, more preferably about 60% pure, even more preferably about 80% pure, most preferably about 90% pure, and even most preferably about 95% pure, as determined by SDS-PAGE. Sometimes such polypeptide is alternatively referred to as a "purified" and/or "isolated" phytase.

20 A phytase polypeptide which comprises a phytase variant of the invention includes fused polypeptides or cleavable fusion polypeptides in which another polypeptide is fused at the N-terminus or the C-terminus of the polypeptide or fragment thereof. A fused polypeptide is produced by fusing a nucleic
25 acid sequence (or a portion thereof) encoding another polypeptide to a nucleic acid sequence (or a portion thereof) encoding a phytase variant of the present invention. Techniques for producing fusion polypeptides are known in the art, and include, ligating the coding sequences encoding the polypeptides
30 so that they are in frame and that expression of the fused

polypeptide is under control of the same promoter(s) and terminator.

A "feed" and a "food," respectively, means any natural or artificial diet, meal or the like or components of such meals
5 intended or suitable for being eaten, taken in, digested, by an animal and a human being, respectively.

The phytase variant of the invention may exert its effect in vitro or in vivo, i.e. before intake or in the stomach of the individual, respectively. Also a combined action is possible.

10 A phytase composition according to the invention always comprises at least one phytase of the invention.

Generally, phytase compositions are liquid or dry.

Liquid compositions need not contain anything more than the phytase enzyme, preferably in a highly purified form.
15 Usually, however, a stabilizer such as glycerol, sorbitol or mono propylen glycol is also added. The liquid composition may also comprise other additives, such as salts, sugars, preservatives, pH-adjusting agents, proteins, phytate (a phytase substrate). Typical liquid compositions are aqueous or oil-based
20 slurries. The liquid compositions can be added to a food or feed after an optional pelleting thereof.

Dry compositions may be spray-dried compositions, in which case the composition need not contain anything more than the enzyme in a dry form. Usually, however, dry compositions are so-
25 called granulates which may readily be mixed with e.g. food or feed components, or more preferably, form a component of a pre-mix. The particle size of the enzyme granulates preferably is compatible with that of the other components of the mixture. This provides a safe and convenient means of incorporating
30 enzymes into e.g. animal feed.

Agglomeration granulates are prepared using agglomeration technique in a high shear mixer (e.g. Lödige) during which a filler material and the enzyme are co-agglomerated to form granules. Absorption granulates are prepared by having cores of
5 a carrier material to absorb/be coated by the enzyme.

Typical filler materials are salts such as disodium sulphate. Other fillers are kaolin, talc, magnesium aluminium silicate and cellulose fibres. Optionally, binders such as dextrans are also included in agglomeration granulates.

10 Typical carrier materials are starch, e.g. in the form of cassava, corn, potato, rice and wheat. Salts may also be used.

Optionally, the granulates are coated with a coating mixture. Such mixture comprises coating agents, preferably hydrophobic coating agents, such as hydrogenated palm oil and
15 beef tallow, and if desired other additives, such as calcium carbonate or kaolin.

Additionally, phytase compositions may contain other substituents such as colouring agents, aroma compounds, stabilizers, vitamins, minerals, other feed or food enhancing
20 enzymes, i.e. enzymes that enhances the nutritional properties of feed/food, etc. This is so in particular for the so-called pre-mixes.

A "food or feed additive" is an essentially pure compound or a multi component composition intended for or suitable for
25 being added to food or feed. In particular it is a substance which by its intended use is becoming a component of a food or feed product or affects any characteristics of a food or feed product. It is composed as indicated for phytase compositions above. A typical additive usually comprises one or more
30 compounds such as vitamins, minerals or feed enhancing enzymes and suitable carriers and/or excipients.

In a preferred embodiment, the phytase compositions of the invention additionally comprises an effective amount of one or more feed enhancing enzymes, in particular feed enhancing enzymes selected from the group consisting of α -galactosidases, β -galactosidases, in particular lactases, other phytases, β -glucanases, in particular endo- β -1,4-glucanases and endo- β -1,3(4)-glucanases, cellulases, xylosidases, galactanases, in particular arabinogalactan endo-1,4- β -galactosidases and arabinogalactan endo-1,3- β -galactosidases, endoglucanases, in particular endo-1,2- β -glucanase, endo-1,3- α -glucanase, and endo-1,3- β -glucanase, pectin degrading enzymes, in particular pectinases, pectinesterases, pectin lyases, polygalacturonases, arabinanases, rhamnogalacturonases, rhamnogalacturonan acetyl esterases, rhamnogalacturonan- α -rhamnosidase, pectate lyases, and α -galacturonisidases, mannanases, β -mannosidases, mannan acetyl esterases, xylan acetyl esterases, proteases, xylanases, arabinoxylanases and lipolytic enzymes such as lipases, phospholipases and cutinases.

The animal feed additive of the invention is supplemented to the mono-gastric animal before or simultaneously with the diet. Preferably, the animal feed additive of the invention is supplemented to the mono-gastric animal simultaneously with the diet. In a more preferred embodiment, the animal feed additive is added to the diet in the form of a granulate or a stabilized liquid.

An effective amount of phytase in food or feed is from about 10-20.000; preferably from about 10 to 15.000, more preferably from about 10 to 10.000, in particular from about 100 to 5.000, especially from about 100 to about 2.000 FYT/kg feed or food.

Examples of other specific uses of the phytase of the invention is in soy processing and in the manufacture of inositol or derivatives thereof.

The invention also relates to a method for reducing 5 phytate levels in animal manure, wherein the animal is fed a feed comprising an effective amount of the phytase of the invention.

Also comprised in this invention is the use of a phytase of the invention during the preparation of food or feed 10 preparations or additives, i.e. the phytase exerts its phytase activity during the manufacture only and is not active in the final food or feed product. This aspect is relevant for instance in dough making and baking.

The invention relates to a phytase variant which, when 15 aligned according to Fig. 1, is amended as compared to a model phytase in at least one of the following positions, using the position numbering corresponding to P_{lycii}:

24; 27; 31; 33; 39; 40; 41; 42; 43; 44; 45; 46; 47; 49; 51; 56;
58; 59; 61; 62; 68; 69; 70; 71; 72; 73; 74; 75; 76; 77; 78; 79;
20 80; 81; 82; 83; 84; 88; 90; 102; 115; 116; 117; 118; 119; 120;
121; 122; 123; 124; 125; 126; 127; 128; 132; 143; 148; 149; 150;
151; 152; 153; 154; 155; 156; 157; 158; 159; 160; 161; 162; 163;
170f; 170g; 171; 172; 173; 184; 185; 186; 187; 187a; 190; 191;
192; 193; 194; 195; 198; 199; 200; 201; 201a; 201b; 201c; 201d;
25 201e; 201f; 202; 203; 203a; 204; 205; 211; 215; 220; 223; 228;
232; 233; 234; 235; 236; 237; 238; 239; 242; 243; 244; 246;
251e; 253; 256; 260; 264; 265; 267; 270; 271; 272; 273; 274;
275; 276; 277; 278; 279; 280; 283; 285; 287; 288; 292; 293; 302;
304; 332; 333; 334; 335; 336; 337; 338; 339; 340; 341; 342; 343;
30 348; 349; 352; 360; 362; 364; 365; 366; 367; 368; 369; 370; 371;

372; 373; 374; 375; 376; 383k; 387; 393; 394; 396; 404; 409;
411; 412; 413; 417; 421; 431.

From these variants we expect amended characteristics, preferably amended activity characteristics. In fact, for
5 several variants such amended characteristics have already been shown (see the experimental part). Like above, "amended" means as compared to the model phytase. "Amended activity characteristics" means amended in at least one phytase activity related respect, such as (non-exclusive list): pH stability,
10 temperature stability, pH profile, temperature profile, specific activity (in particular in relation to pH and temperature), substrate specificity, substrate cleavage pattern, substrate binding, position specificity, the velocity and level of release of phosphate from corn, reaction rate, phytate degradation
15 rate), end level of released phosphate reached.

Preferred amended activity characteristics are amended specific activity, preferably increased, and preferably increased at a pH of 3, 4, 5, or 6; amended pH or temperature profile; and/or amended, preferably increased, thermostability,
20 e.g. of an increased melting temperature as measured using DSC.

Preferred phytase variants are: Phytase variants which, when aligned according to Fig. 1, are amended as compared to a model phytase in at least one of the following positions, using the position numbering corresponding to P_lycii:

25 43; 44; 47; 51; 58; 62; 78; 80; 83; 88; 90; 102; 143; 148; 153;
154; 186; 187a; 195; 198; 201e; 204; 205; 211; 215; 220; 242;
244; 251e; 260; 264; 265; 267; 270; 273; 278; 302; 336; 337;
339; 352; 365; 373; 383k; 404; 417.

The following variants of *A_fumigatus* constitute a
30 subgroup: Q43L; Q270L; G273D,K; N336S; A205E; Y278H; Q43L+Q270L;

Q43L+Q270L+G273D; Q43L+Q270L+G273D+N336S; G273K+A205E;
G273K+A205E+Y278H (see EP 0897010).

Generally, variants of the invention can be deduced or identified as follows: Looking at the alignment according to
5 Fig. 1, comparing two sequences, one of which is a model phytase with improved properties, identifying amino acid differences in relevant positions/areas, and transferring (substituting with) from the model to the other phytase sequence the amino acid in a relevant position.

10 The invention also relates to a process for preparing a phytase variant which includes the above method, and further includes the deducement and synthesis of the corresponding DNA sequence, the transformation of a host cell, the cultivation of the host cell and the recovery of the phytase variant.

15 Relevant positions/areas include those mentioned below in relation to important phytase activity characteristics such as specific activity, thermostability, pH activity/stability.

 The present invention also relates to phytase variants (varied according to a model phytase as defined herein) which
20 are obtainable, preferably obtained, by the process outlined above and which are expected to exhibit an amended characteristic/property, preferably does exhibit such amended characteristic, e.g. an improved specific activity.

 At least the basidiomycete model phytases *P_lycii* and
25 *T_pubescens* exhibit a high specific activity (as determined using the method of Example 2 herein).

 This is an example of a desired property which can be transferred to other phytases, e.g. the other phytases listed in Fig. 1, in particular to the *A_pediades* and the ascomycete
30 phytases such as *A_fumigatus*, *A-ficum*, *consphyA*, by a deducement process such as the one mentioned above.

Thus, amended specific activity, in particular an improved specific activity, in particular at low pH and/or high temperature, is expected from variants, which have been amended in relevant areas, viz. (i) in the amino acid residues which point into the active site cleft; or (ii) in the amino acid residues in the close neighbourhood of these active site residues. Preferably, close neighbourhood means within 10Å from the active site residues.

From the pdb file 1IHP (Brookhaven Database entry of 10 18.03.98 re 1IHP, Structure of Phosphomonoesterase, D.Kostrewa; or as published in Nature Structural Biology, 4, 1997, p. 185-190), active site regions can be identified, using the program INSIGHTII from Molecular Simulations MSI, San Diego, California, and using the subset command, an "active site shell" can be 15 defined comprising those amino acid residues which lie close to the catalytic residues, defined as H59, D339 and R58 in A. ficum phytase (corresponding to Peniophora numbers H71, D335 and R70, respectively). An "active site shell(10Å)" comprises those residues which lie within 10Å from the above catalytic 20 residues.

The residues within 10Å from H71 and D335 are the following (using Peniophora numbers): 41-47, 68-77, 115-118, 120-126, 128, 149-163, 185, 191-193, 199, 243, 270-271, 273-275, 277-279, 288, 332-343, 364-367, 369-375, 394 ("the active site 25 shell(10Å)").

Preferably, a "substrate binding shell" can also be defined which comprises those residues which are in close proximity to the substrate binding site and which can therefore be expected to be in contact with the substrate.

30 This information can be deduced as described above, by docking a sugar analogue to phytin into the active site cleft

(the residues making up the surface of the active site). If a sugar without any phosphate groups is docked into the active site cleft, e.g. alpha-D-glucose (chair conformation, structure provided by the INSIGHTII program), using a fixed distance as shown below, the residues pointing towards the active site cleft can be extracted using the subset command and using a distance of 10Å from the substrate analogue. Alternatively, the compound inositol-1,4,5-triphosphate (Brookhaven database file 1dix. Inositol-1,4,5-triphosphate) can be docked into the active site cleft. This compound and glucose, however, are more or less superimposable.

The distances in Ångström (Å) are: From oxygen atom in position 6 of the alpha-D-glucose to

	atom ND1 of H59:	5.84
15	atom NH2 of R58:	6.77
	atom NH2 of R142:	5.09
	atom ND2 of N340:	3.00
	atom ND1 of H59:	7.76
	atom NH2 of R58:	8.58.

(the Peniophora numbers of the above residues are: H71, R70, R155, N336, H71 and R70, respectively).

In this way, the residues in contact with the substrate are identified as follows (Peniophora numbers): 43-44; 70-80; 83-84; 115; 153; 155-156; 184; 191-192; 198-202; 205; 235; 238; 242; 270; 272-273; 275-277; 332-336; 338; 369; 371 ("the substrate binding shell(10Å)").

Variants being amended in one or more of (1) the active site shell or (2) the substrate binding shell, are strongly expected to have an amended specific activity. This leads to the following joint grouping of positions (still Peniophora numbers and 10Å shells): 41-47, 68-80, 83-84, 115-118, 120-126, 128,

149-163, 184-185, 191-193, 198-201e, 202-203, 205, 235-236, 238-239, 242-243, 270-279, 285, 288, 332-343, 364-367, 369-375, 394.

Preferably, the active site shell and the substrate binding shell are defined as described above using the basidiomycete model phytases of Fig. 1, the Peniophora phytase being a preferred model. A deducement of corresponding variants of other model phytases is possible using the alignment of Fig. 1.

In a preferred embodiment, a distance of 5Å is used in the subset command, thus defining active site and substrate binding shells of a more limited size, e.g. an active site shell comprising the residues 43-44, 69-74, 117, 125, 155-156, 159, 274, 332-340, 370-374 (5Å from H71 and D335), "active site shell(5Å)".

Generally the active site shell and substrate binding shell regions form the basis for selecting random mutagenesis regions. Examples of preferred random mutagenesis regions are regions 69-74, 332-340, 370-374, doping to be added (a 5Å approach); and

regions 57-62, 142-146, 337-343, doping to be added (a 10Å approach).

It is presently contemplated that any amendment in either of these positions will lead to a phytase of amended characteristics, e.g. of an amended specific activity.

The above expression "any amendment in either of the positions" is considered fully equivalent to listing each position and each substitution, e.g. as follows for the above sub-group 41-47:

41A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

42A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

43A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

44A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

45A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

46A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

47A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y.

5 In a preferred embodiment, amended specific activity is expected from the following variants:

42S, G; 43A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y; 45D, S;
47Y, F; 51E, A; 75W, F; 78S, D; 79G; 80K, A; 83I, Q; 84Q, V; 116S;
118V, L; 119E; 120L; 122A; 123N, T; 125S; 126H, S; 127Q, E; 128A, T;
10 151A, S; 152G; 153D, Y; 154Q, D, G; 157V; 158D, A; 159T; 160A, S;
161T, N; 162N; 163W; 184Q, S; 186A, E; 198A, N; 200G, V; 201D;
deletions of one or more of 201a, 201b, 201c, 201d, 201e, 201f -
preferably all; 202S; 205Q, E; 235Y, L; 238L, M; 242P; 270Y, A, L;
271D; 273D, K; 275F, Y; 278T, H; 332F; 336S; 337T, Q; 339V; 340P, A;
15 343A, S; 364W, F; 365V, L; 366D, V; 367K; 368K; 369I, L; 370V; 373S;
374A; 375H; 376M; 393V.

Particularly preferred variants are the following: 78S;
79G; 80A; 83I, Q; 84Q, V; 198A, N; 200G, V; 201D; deletions in one
or more of 201a, 201b, 201c, 201d, 201e, 201f - preferably all
20 deletions; 202S; 205Q, E; 235Y, L; 238L, M; 242P, 273D; 275F, Y.

Other particularly preferred variants are the following:
43A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y; in particular 43M, P;
75W, F; 80K; 153D; 184Q, S; 270Y, A; 332F; 369I, L.

The following variants are especially preferred:
25 43L, G, N, V, A, I, T; 78D; 153Y; 154G; 270L; 273D, K. Double and
triple variants (43L/270L); (43L/270L/273D); (43L/78D) and
(43L/153Y/154G) are also especially preferred. Other preferred
variants are 205E; 278H; 336S.

These especially preferred single, double and triple
30 variants are preferably variants of model phytases which can be

aligned to Fig.1, in particular variants of the specific model phytases listed in Fig. 1.

At least consphyA is known to have a high thermostability. Still further, the thermostability of P_lycii is rather high.

5 This is an example of a desired property which can be transferred to other phytases, e.g. the other phytases listed in Fig. 1, in particular to the basidiomycete phytases such as P_lycii and A_pediades, by a deducement process such as the one mentioned above.

10 Amended thermostability, in particular improved thermostability, is expected on this background from the following variants:

39H,S; 40L,N; 43P; 47Y,F; 49P; 51E,A; 56P; 58D; 61R; 62V;
80K; 83A; 84Y; 172P; 184P; 195T; 198A; 204V; 211L; 223D; 236Y;
15 242P; 246V; 253P; 264R; 265Q; 280A,P; 283P; 287A; 292F,Y; 293A;
302R; 304P; 337S; 348Y; 387P; 396R; 409R; 411K; 412R; 417E;
421F,Y.

The following variants of amended thermostability are particularly preferred: 39S; 40N; 47Y,F; 51A; 83A; 195T; 204V;
20 211L; 242P; 265A.

Further variants of amended thermostability are the following: 42G; 43T,L,G; 44N; 58K,A; 59G; 62I; 69Q; 75F; 78D;
79G; 80A; 81A,G; 82T; 83K,R; 84I; 88I; 90R,A; 102Y; 115N; 118V;
122A; 123Q,N; 125M,S; 126V,S; 127N,Q; 128S,A; 143N,K; 148V,I;
25 154S; 158D; 170fH; 170gA; 171T,N; 172N; 173W, 184S; 186A; 187A;
187aS; 193S; 195V,L; 198V; 201E; 201eT; 202A, 203aT; 204A; 211V,
215P,A; 220L,N; 223H; 228N; 232T; 322E; 235T; 236N; 242S; 244D;
251eQ,E; 256D; 264I; 260A,H; 265A; 267D; 270G; 271D; 273K,D;
278T,H; 287T; 293V; 302H; 337T,G; 338I; 339V,I; 340A; 352K;
30 365A,S; 366S; 367A; 369L; 373S,A; 374S; 376M; 383kE,Q; 404G,A;
411T; 417R; 431E.

Other concepts of the invention, which can be expected to impart an improved thermostability to a phytase, are as follows - considering the 1IHP structure previously referred to and transferring via an alignment according to Fig. 1 as outlined
5 herein:

(A) Introduction of prolin residues in spatial positions where the prolin special dihedral angles are satisfied and the hydrogen bonding network are not hampered and no steric clashes are observed.

10 (B) Filling up holes: By substitution for bigger residues in internal cavities an improvement in stability can often be obtained.

(C) Cystin bridge: Cystin bridges will often make the proteins more rigid and increase the energy of unfolding.

15 Further variants from which amended thermostability is expected according to these concepts of (A) to (C) are: 27P, 31Y, 132F, 132I, 132L, 184P, 186P, 190P, 280P, 343F, 343I, 343L, 349P, 362P and (33C and 24C).

Concept (A): 27P, 184P, 186P, 190P, 349P, 362P.

20 Concept (B): 343F,I,L; 31Y; 132F,I,L; 273F.

Concept (C): 33C/24C.

Amended pH activity or stability, preferably stability, in particular at low pH, in particular improved, is another desired property which can be transferred by aligning according to Fig.
25 1 and transferring from models of improved pH profiles to other phytases - as outlined above.

Other concepts of the invention, which can be expected to impart an improved stability at low pH to a phytase, are as follows - considering the 1IHP structure previously referred to
30 and transferring via an alignment according to Fig. 1 as outlined herein:

(D) Surface charges: Better distribution at low pH, to avoid cluster of negative or positive, and to avoid too close same charged residues.

(E) Prevent deamidation: Surface exposed Q or N in close
5 contact to negative charged residues.

Phytase variants having improved pH stability/activity at low pH are expected to be: 39H; 39Q; 80A; 203R; 271N; 51R; 154S; 185S; 194S; 194T; 288L; 288I; 288F; 360R; 173Q,S; 204Q,S; 303K,S; 81Q,E.

10 Concept (D): 203R, 271N, 51R, 185S, 360R; 173Q,S; 204Q,S; 303K,S; 81Q,E.

Concept (E): 154S; 194S,T; 288L,I,F.

A preferred model phytase for these concepts of (D) and
(E) is P_lycii.

15 Experimentally proven to have a lowered pH optimum is: Variant 80A of ascomycete phytases, in particular of A_fumigatus and consphyA.

Especially preferred single, double and triple variants are 43L; (43L/270L) and (43L/270L/273D). These variants have a
20 changed pH profile. They are preferably variants of the specific model phytases listed in Fig. 1.

For all preferred variants listed above:

the stability is preferably amended at high temperature, viz. in the temperature range of 50-100°C, in particular 60-
25 90°C, more preferably in the range of 70-90°C;

the activity is preferably amended in a temperature range relevant for the use in the gastro-intestinal system of animals, e.g. 30-40°C, more preferably 32-38°C, most preferably in the range of 35-38°C;

30 the stability is preferably amended at low pH, viz. in the pH range of pH 1.5-7, preferably 2-6, more preferably 3-5;

the activity is preferably amended in the pH range of pH 1.5-5.5, more preferably at pH 2.5-4.5, still more preferably 3-5.

Tests for amended phytase characteristics, such as those 5 mentioned above, are well known in the art and any such test can be used to compare the performance of the phytase variants with the phytase models.

A preferred test for specific activity is given in Example 2. Preferred tests for pH and temperature activity and stability 10 are given in Example 3. An even more preferred test for thermal stability is the DSC method of Example 4.

WO 98/28409 discloses tests for various other parameters, too, such as position specificity. All the tests of WO 98/28409 are preferred tests.

15 Generally, of course all these tests can be conducted at desired pH values and temperatures.

In the dependent claims, some preferred phytase variants based on five of the thirteen herein specifically disclosed model phytases are specified.

20 In an analogous way other preferred variants based on the remaining eight specifically disclosed model phytases can easily be deduced by combining the suggested amendments with each of the corresponding sequences of Fig. 1. These preferred variants are specifically included in the present invention, and they are 25 easily deduced, viz. the following:

Variants of a model phytase derived from Paxillus, preferably Paxillus involutus, preferably derived from strain CBS 100231, preferably variants of P_involutus-A1, the sequence of which is shown at Fig. 2, said variants comprising at least 30 one of the following amendments:

()24C; T27P; F31Y; I33C; R39H,S,Q; N40L; S42G;

P43A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y; Y44N; S45D;
 Y47F; A51E,R; A58D,K; Q61R; I62V; F75W; S78D; A80K; T81Q,E,G,A;
 R83A,I,Q,K; I84Y,Q,V; L88I; K90R,A; F102Y; S115N; D116S; V118L;
 P119E; F120L; A123N,T,Q; S125M; F126H,S,V; D127Q,E,N; A128T,S;
 5 A132F,I,L; I148V; D151A,S; S153D,Y; D154Q,S,G; D158A; S159T;
 A160S; T161N; ()170fH; ()170gA; S171N; H172P; N173Q,S; P184Q,S;
 Q185S; T186A,E,P; G187A; ()187aS; T190P,A; D193S; N194S,T;
 M195T,V,L; A198N,V; G200V; D201E; ()201eT; S202A; D203R,K,S;
 P203aV,T; Q204E,S,A,V; V205E; V211L; S215A,I; L220N; A223D,H;
 10 D233E; F235Y,L,T; N236Y; L237F; V238L,M; A242P,S; M244D;
 ()251eE,Q; D253P; T256D; P260A,H; E264R,I; A265Q; A267D;
 G270Y,A,L; D271N; D273K; F275Y; T278H; Y280A,P; E283P; V287A,T;
 Q288L,I,F; Y292F; V293A; N302R,H; A304P; N336S; L337T,Q,S,G; M
 338I; V339I; A340P; S343A,F,I,L; F348Y; R349P; A352K; P360R;
 15 R362P; W364F; R365V,L,A,S; T366D,V,S; S367K,A; S368K; L369I;
 S373A; G374A,S; R375H; ()383kQ,E; T387P; Q396R; G404A; L409R;
 T411K; L412R; E417R; F421Y.

Variants of a model phytase derived from a species of the
 genus *Paxillus*, preferably the species *Paxillus involutus*,
 20 preferably derived from strain CBS 100231, preferably variants
 of *P_involutus*-A2, the sequence of which is shown at Fig. 3, said
 variants comprising at least one of the following amendments:

P24C; I27P; F31Y; I33C; R39H,S,Q; N40L; S42G;

P43A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y; Y44N; S45D;
 25 Y47F; A51E,R; A58D,K; E61R; I62V; F75W; S78D; A80K; A81Q,E,G;
 R83A,I,Q,R,K; I84Y,Q,V; L88I; K90R,A; F102Y; S115N; D116S;
 V118L; P119E; F120L; A123N,T,Q; S125M; F126H,S,V; D127Q,E,N;
 A128T,S; V132F,I,L; D143N; I148V; D151A,S; S153D,Y; D154Q,S,G;
 D158A; A160S; T161N; ()170fH; ()170gA; S171N; R172P; N173Q,S;
 30 P184Q,S; Q185S; T186A,E,P; G187A; ()187aS; T190P,A; D193S;
 N194S,T; M195T,V,L; A198N,V; G200V; E201D; ()201eT; S202A;

D203R,K,S; P203aV,T; Q204E,S,A,V; V205E; S211L,V; S215A,P;
 L220N; A223D,H; A232T; F235Y,L,T; N236Y; L237F; V238L,M; P242S;
 M244D; ()251eE,Q; D253P; T256D; P260A,H; E264R,I; A265Q; A267D;
 G270Y,A,L; D271N; D273K; F275Y; T278H; Y280A,P; A283P; V287A,T;
 5 Q288L,I,F; Y292F; I293A,V; N302R,H; A304P; N336S; L337T,Q,S,G;
 M338I; V339I; 340P,A; A343S,F,I,L; F348Y; R349P; A352K; P360R;
 R362P; W364F; L365V,A,S; T366D,V,S; S367K,A; S368K; V369I,L;
 S373A; R375H; ()383kQ,E; T387P; Q396R; G404A; L409R; A411K,T;
 L412R; E417R; Y421F.

10 Variants of a model phytase derived from a species of the
 genus *Trametes*, preferably the species *Trametes pubescens*,
 preferably derived from strain CBS 100232, preferably variants
 of *T. pubescens*, the sequence of which is shown at Fig. 4, said
 variants comprising at least one of the following amendments:

15 R24C; T27P; L31Y; V33C; Q39H,S; S40L,N; S42G;
 M43A,C,D,E,F,G,H,I,K,L,N,P,Q,R,S,T,V,W,Y; Y44N; S45D;
 Y47F; A51E,R; A58D,K; S59G; Q61R; I62V; F75W; S78D; A80K;
 A81Q,E,G; R83A,I,Q,K; I84Y,Q,V; V88I; K90R,A; L102Y; D115N;
 V118L; T123N,Q; S125M; S126H,V; E127Q,N; A128T,S; A132F,I,L;
 20 D143N; V148I; S151A; S153D,Y; D154Q,S,G; A158D; A160S; N161T;
 ()170fH; ()170gA; S171N; S172P; N173Q,S; S184Q,P; E185S;
 A186E,P; G187A; ()187aS; T190P,A; N194S,T; M195T,V,L; A198N,V;
 G200V; ()201eT; S202A; D203R,K,S; P203aV,T; Q204E,S,A,V; V205E;
 Q211L,V; P215A; L220N; G223D,H; D233E; Y235L,T; N236Y; L237F;
 25 L238M; P242S; E244D; ()251eE,Q; E253P; Q260A,H; D264R,I; A265Q;
 A267D; A270Y,L,G; D271N; D273K; F275Y; T278H; Y280A,P; V287A,T;
 Q288L,I,F; Y292F; I293A,V; A302R,H; N304P,A; N336S; Q337T,S,G;
 M338I; V339I; A340P; S343A,F,I,L; F348Y; N349P; A352K; P360R;
 R362P; F364W; L365V,A,S; V366D,S; K367A; I369L; A373S; A374S;
 30 R375H; ()383kQ,E; Q387P; A396R; G404A; V409R; T411K; L412R;
 E417R; Y421F.

Variants of a model phytase derived from a species of the genus *Aspergillus*, preferably the species *Aspergillus nidulans*, preferably derived from strain DSM 9743, preferably variants of *A_nidulans*, the sequence of which is shown at Fig. 10, said

5 variants comprising at least one of the following amendments:

V24C; A27P; H39S,Q; V40L,N; G42S;

Q43A,C,D,E,F,G,H,I,K,L,M,N,P,R,S,T,V,W,Y; Y44N; S45D;
 Y47F; S49P; E51A,R; V56P; H58D,K,A; E61R; V62I; S69Q; Y75W,F;
 E78D,S; S79G; K80A; S81Q,E,A,G; K82T; A83I,Q,K,R; Y84Q,V,I;
 10 A90R; D115N; D116S; T118V,L; I119E; F120L; E122A; N123T,Q;
 M125S; V126H,S; D127Q,E,N; S128A,T; F132I,L; K143N; I148V;
 S151A; S153D,Y; D154Q,S,G; A158D; S159T; A160S; E161T,N; K162N;
 F163W; G170fH; S170gA; ()171N; ()172P; K173Q,S; P184Q,S; E185S;
 I186A,E,P; D187A; G187aS; T190P,A; H193S; S194T; S198A,N,V;
 15 E200G,V; N201D,E; D201e(); E201e(),T; R201f() (a deletion of at
 least one of 201d, 201e, 201f, preferably all); A202S;
 D203R,K,S; E203aV,T; I204Q,E,S,A,V; I211L,V; P215A; L220N;
 D223H; K228N; E232T; N233E; I235Y,L,T; Y236N; L237F; M238L;
 S242P; M246V; E251eQ; A256D; E260A,H; L264R,I; Q270Y,A,L,G;
 20 S271D,N; S273D,K; Y275F; G278T,H; A280P; A287T; Q288L,I,F;
 F292Y; T293A,V; Q302R,H; P304A; N336S; S337T,Q,G; M338I; I339V;
 S340P,A; F343A,S,I,L; N349P; Q352K; S360R; Q362P; Y364W,F;
 A365V,L,S; A366D,V,S; S367K,A; W368K; T369I,L; G373S,A; A374S;
 R375H; A376M; E383kQ; A404G; T411K; L412R; E417R; F421Y; K431E.

25 Variants of a model phytase derived from a species of *Aspergillus*, preferably *Aspergillus terreus*, preferably derived from strain CBS 220.95, preferably variants of *A_terreus*, the sequence of which is shown at Fig. 12, said variants comprising at least one of the following amendments:

30 G24C; V27P; H39S,Q; K40L,N; G42S;

L43A,C,D,E,F,G,H,I,K,M,N,P,Q,R,S,T,V,W,Y; Y44N; A45D,S;
 Y47F; S49P; Q51E,A,R; V56P; P58D,K,A; D59G; H61R; I62V; A69Q;
 S75W,F; H78D,S; S79G; K80A; T81Q,E,A,G; A83I,Q,K,R; Y84Q,V,I;
 A90R; E115N; E116S; T118V,L; P119E; F120L; R122A; N123T,Q;
 5 L125S,H; R126H,S,V; D127Q,E,N; L128A,T,S; F132I,L; H143N; V148I;
 T151A,S; D152G; A153D,Y; S154D,Q,G; H157V; E158D,A; S159T;
 A160S; E161T,N; K162N; F163W; H173Q,S; P184Q,S; E185S;
 G186A,E,P; S187A; A187aS; T190P,A; H193S; S194T; L195T,V;
 A198N,V; E200G,V; S201D,E; S201d(); T201e(); V201f(); G202S,A;
 10 D203R,K,S; D203aV,T; A204Q,E,S,V; V205E; V211L; A215P; L220N;
 D223H; Q228N; D232T; D233E; V235Y,L,T; N236Y; L237F; M238L;
 P242S; E244E; T251eE,Q; A260H; T264R,I; Q265A; N267D; L270Y,A,G;
 S271D,N; K273D; Y275F; H278T; G280A,P; V287A,T; Q288L,I,F;
 W292F,Y; A293V; Q302H; P304A; N337T,Q,S,G; L338I; V339I;
 15 S340P,A; W343A,S,F,I,L; N349P; A352K; S360R; S362P; Y364W,F;
 A365V,L,S; A366D,V,S; A367K; W368K; T369I,L; A373S; A374S;
 R375H; A376M; R383kQ,E; P404A,G; K411T; A417E,R; F421Y; A431E.

Variants of a model phytase derived from a species of
Talaromyces, preferably the species *Talaromyces thermophilus*,
 20 preferably derived from strain ATCC 20186 or ATCC 74338,
 preferably variants of *T_thermo*, the sequence of which is shown
 at Fig. 13, said variants comprising at least one of the
 following amendments:

H24C; V27P; H39S,Q; S40L,N; G42S;
 25 Q43A,C,D,E,F,G,H,I,K,L,M,N,P,R,S,T,V,W,Y; Y44N; S45D;
 F47Y; S49P; A51E,R; V56P; Q58D,K,A; N59G; K61R; I62V; Y75W,F;
 S78D; S79G; K80A; T81Q,E,A,G; E82T; L83A,I,Q,R,K; Y84Q,V,I;
 R90A; D116S; T118V,L; P119E; F120L; E122A; N123T,Q; M125S;
 I126H,S,V; Q127E,N; L128A,T,S; F132I,L; V148I; S151A; S153D,Y;
 30 D154Q,S,G; I157V; A158D; S159T; G160A,S; R161T,N; L162N; F163W;
 S170gA; D171N; K172P; H173Q,S; E184Q,S,P; E185S; G186A,E,P;

D187A; T190P,A; T193S; G194S,T; S195T,V,L; V198A,N; E200G,V;
 D201E; S201d(); S201e(),T; S201f(); G202S,A; H203R,K,S;
 D203aV,T; A204Q,E,S,V; Q205E; Q211L,V; A215P; I220N,L; H223D;
 D228N; S232T; D233E; P235Y,L,T; Y236N; M237F; D238L,M; P242S;
 5 E244D; L246V; ()251eE,Q; A256D; Q260A,H; Q264R,I; A265Q;
 Q270Y,A,L,G; S271D,N; G273D,K; Y275F; N278T,H; G280A,P; A287T;
 Q288L,I,F; F292Y; V293A; H302R; P304A; N336S; T337Q,S,G; M338I;
 T339V,I; S340P,A; A343S,F,I,L; N349P; A352K; S360R; E362P;
 Y364W,F; S365V,L,A; A366D,V,S; A367K; W368K; T369I,L; G373S,A;
 10 G374A,S; R375H; A376M; D383kQ,E; E404A; K411T; R417E; F421Y.

Variants of a model phytase derived from a species of
 Thermomyces, preferably the species Thermomyces lanuginosus,
 preferably derived from strain DBS 586.94, preferably variants
 of T_lanuginosa, the sequence of which is shown at Fig. 14, said
 15 variants comprising at least one of the following amendments:

K24C; ()27P; ()31Y; ()33C; R39H,S,Q; H40L,N; G42S;
 Q43A,C,D,E,F,G,H,I,K,L,M,N,P,R,S,T,V,W,Y; Y44N; S45D;
 F47Y; S49P; A51E,R; V56P; K58D,A; V62I; S69Q; Y75W,F; A78D,S;
 H79G; K80A; S81Q,E,A,G; E82T; V83A,I,Q,K,R; Y84Q,V,I; L88I;
 20 R90A; F102Y; D115N; N116S; T118V,L; R119E; F120L; E122A;
 E123N,T,Q; M125S; M126H,S,V; E127Q,N; S128A,T; F132I,L; E143N;
 V148I; A151S; S153D,Y; A154D,Q,S,G; I157V; A158D; S159T; A160S;
 E161T,N; F162N; F163W; R170fH; S170gA; K172P; D173Q,S; S184Q,P;
 E185S; E186A,P; T187A; G187aS; T190P,A; G193S; L194S,T; T195V,L;
 25 A198N,V; E200G,V; E201D; A201d(); P201e(),T; D202S,A; P203R,K,S;
 T203aV; Q204E,S,A,V; P205E; V211L; R215A,P; I220L,N; H223D;
 E232T; D233E; P235Y,L,T; L236Y,N; M238L; P242S; Q251eE; H256D;
 Q260H; M264R,I; A265Q; Y270A,L,G; T271D,N; D273K; Y275F; H278T;
 G280A,P; A283P; S287A; R288L,I,F; F292Y; V293A; G302R,H; P304A;
 30 N336S; T337Q,S,G; M338I; T339V,I; G340P,A; S343A,F,I,L; N349P;
 P360R; T362P; Y364W,F; A365V,L,S; A366D,V,S; S367K,A; W368K;

T369I,L; A373S; A374S; R375H; A376M; E383kQ; R404A,G; R411K,T;
K417E,R; F421Y; D431E.

Variants of a model phytase derived from a species of
Myceliophthora, preferably the species Myceliophthora
5 thermophila, preferably derived from strain ATCC 48102 or ATCC
74340, preferably variants of M_thermophila, the sequence of
which is shown at Fig. 7, said variants comprising at least one
of the following amendments:

S24C; F31Y; H39S,Q; F40L,N; G42S;
10 Q43A,C,D,E,F,G,H,I,K,L,M,N,P,R,S,T,V,W,Y; Y44N; S45D;
Y47F; S49P; P51E,A,R; I56P; D58K,A; D59G; E61R; V62I; S69Q;
A75W,F; L78D,S; K79G; R80K,A; A81Q,E,G; A82T; S83A,I,Q,K,R;
Y84Q,V,I; R90A; D115N; E116S; T118V,L; R119E; T120L; Q122A;
Q123N,T; M125S; V126H,S; N127Q,E; S128A,T; F132I,L; K143N;
15 V148I; A151S; Q153D,Y; D154Q,S,G; H158D,A; S159T; A160S;
E161T,N; G170fH; S170gA; T171N; F163W; V172P; R173Q,S; P184Q,S;
E185S; T186A,E,P; G187aS; T190P,A; N193S; D194S,T; L195T,V;
A198N,V; E200G,V; E201D; G201a(); P201b(); Y201c(); S201d();
T201e(); I201f(); G202S,A; D203R,K,S; D203aV,T; A204Q,E,S,V;
20 Q205E; T211L,V; P215A; V220N,L; N223D,H; A232T; D233E;
V235Y,L,T; A236Y,N; L237F; M238L; P242S; E244D; A251eE,Q; R256D;
E260A,H; R264I; A265Q; Q270Y,A,L,G; S271D,N; K273D; Y275F;
Y278T,H; P280A; T287A; Q288L,I,F; F292Y; V293A; ()302R,H; P304A;
N336S; D337T,Q,S,G; M338I; M339V,I; G340P,A; G343A,S,F,I,L;
25 D349P; P352K; D360R; E362P; Y364W,F; A365V,L,S; A366D,V,S;
S367K,A; W368K; A369I,L; A373S; A374S; R375H; I376M; E383kQ;
E387P; G404A; M409R; T411K; L412R; E417R; F421Y; D431E.

This invention also provides a new phytase which has been
derived from a strain of Cladorrhinum, viz. C. foecundissimum.
30 Accordingly, the invention also relates to a polypeptide having
phytase acitivity and which comprises SEQ ID NO:2 or the mature

part (amino acids nos 16-495) thereof; or a polypeptide being at least 70, more preferably 75, 80, 85, 90, 95% homologous thereto; homology meaning similarity, preferably identity, and being determined using the program GAP and the settings as defined hereinabove. And the invention relates to a DNA construct which encodes a polypeptide having phytase activity, said DNA construct comprising a DNA molecule which comprises SEQ ID NO:1 or nucleotides nos. 20-70 and 207-1560 thereof; or nucleotides nos. 20-70 and 207-1563 thereof; or nucleotides nos. 65-70 and 207-1563 thereof; or a DNA construct or molecule which is at least 70, 75, 80, 85, 90, 95 % homologous to either of these nucleotide sequences; homology meaning similarity, preferably identity, and being determined using computer programs known in the art such as GAP provided in the GCG program package (Program Manual for the Wisconsin Package, Version 8, August 1996, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711) (Needleman, S.B. and Wunsch, C.D., (1970), Journal of Molecular Biology, 48, 443-453). Using GAP with the following settings for DNA sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3. The invention also relates to a DNA construct which hybridizes with any of the above DNA sequences under the conditions mentioned hereinabove.

25 EXAMPLES

Example 1

Phytase activity assay (FYT)

Phytase activity can be measured using the following assay:

30 10 µl diluted enzyme samples (diluted in 0.1 M sodium acetate, 0.01 % Tween20, pH 5.5) are added into 250 µl 5 mM sodium

phytate (Sigma) in 0.1 M sodium acetate, 0.01 % Tween20, pH 5.5 (pH adjusted after dissolving the sodium phytate; the substrate is preheated) and incubated for 30 minutes at 37°C. The reaction is stopped by adding 250 µl 10 % TCA and free phosphate is
5 measured by adding 500 µl 7.3 g FeSO₄ in 100 ml molybdate reagent (2.5 g (NH₄)₆Mo₇O₂₄·4H₂O in 8 ml H₂SO₄ diluted to 250 ml). The absorbance at 750 nm is measured on 200 µl samples in 96 well microtiter plates. Substrate and enzyme blanks are included. A phosphate standard curve is also included (0-2 mM
10 phosphate). 1 FYT equals the amount of enzyme that releases 1 µmol phosphate/min at the given conditions.

Example 2

Test for specific activity

15 The specific activity can be determined as follows:

A highly purified sample of the phytase is used (the purity is checked beforehand on an SDS poly acryl amide gel showing the presence of only one component).

The protein concentration in the phytase sample is
20 determined by amino acid analysis as follows: An aliquot of the phytase sample is hydrolyzed in 6N HCl, 0.1% phenol for 16 h at 110 C in an evacuated glass tube. The resulting amino acids are quantified using an Applied Biosystems 420A amino acid analysis system operated according to the manufacturers instructions.
25 From the amounts of the amino acids the total mass - and thus also the concentration - of protein in the hydrolyzed aliquot can be calculated.

The activity is determined in the units of FYT. One FYT equals the amount of enzyme that liberates 1 micromol inorganic
30 phosphate from phytate (5 mM phytate) per minute at pH 5.5, 37°C; assay described e.g. in example 1.

The specific activity is the value of FYT/mg enzyme protein.

Example 3

5 Test for temperature and pH activity and stability

Temperature and pH activity and stability can be determined as follows:

Temperature profiles (i.e. temperature activity relationship) by running the FYT assay of Example 1 at various 10 temperatures (preheating the substrate).

Temperature stability by pre-incubating the phytase in 0.1 M sodium phosphate, pH 5.5 at various temperatures before measuring the residual activity.

The pH-stability by incubating the enzyme at pH 3 (25 mM 15 glycine-HCl), pH 4-5 (25 mM sodium acetate), pH 6 (25 mM MES), pH 7-9 (25 mM Tris-HCl) for 1 hour at 40°C, before measuring the residual activity.

The pH-profiles (i.e. pH activity relationship) by running the assay at the various pH using the same buffer-systems (50 20 mM, pH re-adjusted when dissolving the substrate).

Example 4

DSC as a preferred test for thermostability

The thermostability or melting temperature, T_m , can be 25 determined as follows:

In DSC the heat consumed to keep a constant temperature increase in the sample-cell is measured relative to a reference cell. A constant heating rate is kept (e.g. 90°C/hour). An endo-thermal process (heat consuming process - e.g. the unfolding of 30 an enzyme/protein) is observed as an increase in the heat

transferred to the cell in order to keep the constant temperature increase.

DSC can be performed using the MC2-apparatus from MicroCal. Cells are equilibrated 20 minutes at 20°C before scanning to 90°C at a scan rate of 90°/h. Samples of e.g. around 2.5 mg/ml phytase in 0.1 M sodium acetate, pH 5.5 are loaded.

Example 5

Phytase variants of amended activity characteristics

Variants of an *Aspergillus fumigatus* model phytase (a wild type phytase derived from strain ATCC 13073) were prepared as described in EP 98104858.0 (EP-A-0897010), examples 2-3 and 5, and the phytase activity was determined as described in example 7 thereof. pH- and temperature optimum and melting point was determined as described in examples 9 and 10 of EP 98113176.6 (EP-A-0897985).

In Table 1, variants of improved specific activity at pH 5.0 are listed. Table 2 lists variants of improved relative activity at pH 3.0, and Table 3 lists variants of improved thermostability (temperature optimum, e.g. determined by DSC).

Table 1

Amended in position no.	Substitution into	Specific activity at pH 5.0 (U/mg)
43	43L	83.4
	43N	45.5
	43T	106.9
	43I	91.2
	43V	35.0
	43A	27.3
	43G	59.6

43 and 270	43L, 270L	88.7
43 and 270 and 273	43L, 270L, 273D	92.3
43 and 78	43L, 78D	118.5
43 and 153 and 154	43L, 153Y, 154G	193.0
A. fumigatus wild-type phytase	-	26.5

Table 2

Amended in position no.	Substitution into	Relative phytase activity at pH 3.0
205	205E	41%
273	273K	61%
278	278H	75%
273 and 205	273K, 205E	65%
273 and 278	273K, 278H	100%
273 and 205 and 278	273K, 205E, 278H	96%
A. fumigatus wild-type phytase	-	32%

Table 3

Amended in position no.	Substitution into	Temperature optimum (°C)	Tm (°C) (DSC)
43 and 47 and 88 and 102 and 220 and 242 and 267	43T, 47Y, 88I, 102Y, 220L, 242P, 267D	60	67
as above plus 51 and 302 and 337 and 373 and 115	as above plus 51A, 302H, 337T, 373A, 115N	63	-

A. fumigatus wild-type phytase	-	55	62.5
--------------------------------	---	----	------

Example 6**Further phytase variants of amended activity characteristics**

Variants of the ascomycete consensus sequence "conphys" of Fig. 9 were prepared as described in EP 98113176.6 (EP-A-5 0897985), examples 4-8. Phytase activity, including pH- and temperature optimum, and melting point was determined as described in examples 9 and 10, respectively, thereof.

The tables below list variants of amended activity characteristics, viz.

10 Table 4 variants of improved specific activity at pH 6.0;

Table 5 variants of amended pH optimum (the pH-optimum indicated is an approximate value, determined as that pH-value (selected from the group consisting of pH 4.0; 4.5; 5.0; 5.5; 6.0; 6.5; and 7.0) at which the maximum phytase activity was
15 obtained);

Table 6 a variant of improved thermostability (expressed by way of the melting point as determined by differential scanning calorimetry (DSC)); and

20 Table 7 variants of amended thermostability (temperature optimum); a "+" or "-" indicates a positive or a negative, respectively, effect on temperature optimum of up to 1°C; and a "++" and "--" means a positive or a negative, respectively, effect on temperature optimum of between 1 and 3°C.

25 Table 4

Amended in position no.	Substitution into	Specific activity at pH 6.0 (U/mg)
43	43T	130

	43L	205
Conphys	-	62

Table 5

Amended in position no.	Substitution into	pH optimum around
43	43T	6.0
	43L	5.5
	43G	6.5
43 and 44	43L, 44N	6.0
	43T, 44N	5.5
Conphys	-	6.0

Table 6

Amended in position no.	Substitution into	T _m (°C)
43	43T	78.9
Conphys	-	78.1

5 Table 7

Amended in position no.	Substitution into	Temperature optimum amendment
51	A	+
58	K	+
220	N	+
195	L	++
201e	T	++
244	D	+
264	I	+
302	H	+

337	T	++
352	K	+
373	A	++
47	F	-
62	I	-
83	K	-
90	R	-
143	N	-
148	V	---
186	A	---
187a	S	-
198	V	-
204	A	---
211	V	-
215	P	---
251e	Q	-
260	A	-
265	A	-
339	V	-
365	A	---
383k	E	-
404	G	---
417	R	---
Conphys	-	0

Table 8

Amended in position no.	Substitution into	T _m (°C) (DSC)	Specific activity at pH 5.0 (U/mg)
43 and 51 and 220 and 244 and 264 and 302 and 337 and 352 and 373	51A, 220N, 244D, 264I, 302H, 337T, 352K, 373A, 43T	84.7	105
as above plus 80	as above plus 80A	85.7	180
Conphys	-	78.1	30

Example 7**Cloning of a phytase of *Cladorrhinum foecundissimum***

DNA encoding a phytase from *Cladorrhinum foecundissimum* CBS 427.97 has been cloned, and the enzyme isolated and purified, essentially as described in WO 98/28409.

Fig. 15 shows the DNA sequence of the HindIII/XbaI cloned PCR product in pA2phy8. The cloned PCR product is amplified from the genomic region encoding *Cladorrhinum foecundissimum* CBS 427.97 phyA gene. The putative intron is indicated by double underline of the excision-ligation points in accordance with the GT-AG rule (R. Breathnach et al. Proc. Natl. Acad. Sci. USA 75 (1978) pp4853-4857). The restrictions sites used for cloning are underlined.

According to the SignalP V1.1 prediction (Henrik Nielsen, Jacob Engelbrecht, Stren Brunak and Gunnar von Heijne: 5 "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites," Protein Engineering 10, 1-6 (1997)), the signal peptide part of the enzyme corresponds

to amino acids nos. 1-15, accordingly the mature enzyme is amino acids nos. 16-495.

The enzyme exhibits a pH optimum around pH 6 with no activity at the low pH (pH 3), but significant activity up until 5 pH 7.5; thus it is a more alkaline phytase as compared to the *Aspergillus ficuum* phytase.

A temperature optimum around 60°C was found at pH 5.5. Thus, this phytase is more thermostable than the *A. ficuum* phytase.

10

Example 8

Alignment of a new model phytase according to Fig. 1

The phytase sequence of *Cladorrhinum foecundissimum* as disclosed in Example 7 is compared with the 13 model phytases of 15 Fig. 1 using GAP version 8 referred to above with a GAP weight of 3.000 and a GAP lengthweight of 0.100. Complete amino acid sequences are compared. The *M_thermophila* phytase sequence turns up to be the most homologous sequence, showing a degree of similarity to the *C. foecundissimum* sequence of 70.86%.

20 Still using the GAP program and the parameters mentioned above, the phytase sequence "C_foecundissimum" is now aligned to the "M-thermophila" phytase - see Fig. 16. The average match is 0.540; the average mismatch -0.396; quality 445.2; length 505; ratio 0.914; gaps 9; percent similarity 70.860; percent identity 25 53.878.

In a next step, see Fig. 17, the *C_foecundissimum* is pasted (or it could simply be written) onto the alignment of Fig. 1 as the bottom row, ensuring that those amino acid residues which according to the alignment at Fig. 16 are 30 identical (indicated by a vertical line) or similar (indicated by one or two dots) are placed above each other. At 5 places along the sequence, the *C_foecundissimum* sequence comprises

"excess" amino acid residues, which the alignment of Fig. 1 does not make room for. At Fig. 17, these excess residues are transferred onto a next row (but they can be included in the multiple alignment and numbered as described previously in the position numbering related paragraphs (using the denotations a, b, c etc.)).

Corresponding variants of the phytase of *C_foecundissimum* are then easily deduced on the basis of Fig. 17. Some examples: The variants generally designated "80K,A" and "43T" in *C_foecundissimum* correspond to "K80A" and "Q43T," respectively.

CLAIMS

1. A phytase variant which, when aligned according to Fig. 1, is amended as compared to a model phytase in at least one of the following positions, using the position numbering corresponding to P_lycii:

24; 27; 31; 33; 39; 40; 41; 42; 43; 44; 45; 46; 47; 49; 51; 56;
58; 59; 61; 62; 68; 69; 70; 71; 72; 73; 74; 75; 76; 77; 78; 79;
80; 81; 82; 83; 84; 88; 90; 102; 115; 116; 117; 118; 119; 120;
121; 122; 123; 124; 125; 126; 127; 128; 132; 143; 148; 149; 150;
10 151; 152; 153; 154; 155; 156; 157; 158; 159; 160; 161; 162; 163;
170f; 170g; 171; 172; 173; 184; 185; 186; 187; 187a; 190; 191;
192; 193; 194; 195; 198; 199; 200; 201; 201a; 201b; 201c; 201d;
201e; 201f; 202; 203; 203a; 204; 205; 211; 215; 220; 223; 228;
232; 233; 234; 235; 236; 237; 238; 239; 242; 243; 244; 246;
15 251e; 253; 256; 260; 264; 265; 267; 270; 271; 272; 273; 274;
275; 276; 277; 278; 279; 280; 283; 285; 287; 288; 292; 293; 302;
304; 332; 333; 334; 335; 336; 337; 338; 339; 340; 341; 342; 343;
348; 349; 352; 360; 362; 364; 365; 366; 367; 368; 369; 370; 371;
372; 373; 374; 375; 376; 383k; 387; 393; 394; 396; 404; 409;
20 411; 412; 413; 417; 421; 431.

2. A phytase variant which, when aligned according to Fig. 1, comprises at least one of the following amendments as compared to a model phytase, using the position numbering corresponding to the phytase of P_lycii:

24C; 27P; 31Y; 33C; 39H,S,Q; 40L,N; 42S,G;
43A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y; 44N; 45D,S; 47Y,F;
49P; 51E,A,R; 56P; 58D,K,A; 59G; 61R; 62V,I; 69Q; 75W,F; 78D,S;
79G; 80K,A; 81A,G,Q,E; 82T; 83A,I,K,R,Q; 84I,Y,Q,V; 88I; 90R,A;
30 102Y; 115N; 116S; 118V,L; 119E; 120L; 122A; 123N,Q,T; 125M,S;

126H,S,V; 127Q,E,N; 128A,S,T; 132F,I,L; 143N; 148V,I; 151A,S;
152G; 153D,Y; 154D,Q,S,G; 157V; 158D,A; 159T; 160A,S; 161T,N;
162N; 163W; 170fH; 170gA; 171N; 172P; 173Q,S; 184Q,S,P; 185S;
186A,E,P; 187A; 187aS; 190A,P; 193S; 194S,T; 195T,V,L; 198A,N,V;
5 200G,V; 201D,E; 201a(); 201b(); 201c(); 201d(); 201e(); 201f();
201eT; 202S,A; 203R,K,S; 203aV,T; 204Q,E,S,A,V; 205E; 211L,V;
215A,P; 220L,N; 223H,D; 228N; 232T; 233E; 235Y,L,T; 236Y,N;
237F; 238L,M; 242P,S; 244D; 246V; 251eE,Q; 253P; 256D; 260A,H;
264R,I; 265A,Q; 267D; 270Y,A,L,G; 271D,N; 273D,K; 275F,Y;
10 278T,H; 280A,P; 283P; 287A,T; 288L,I,F; 292F,Y; 293A,V; 302R,H;
304P,A; 332F; 336S; 337T,G,Q,S; 338I; 339V,I; 340P,A;
343A,S,F,I,L; 348Y; 349P; 352K; 360R; 362P; 364W,F; 365V,L,A,S;
366D,S,V; 367A,K; 368K; 369I,L; 370V; 373A,S; 374S,A; 375H;
376M; 383kQ,E; 387P; 393V; 396R; 404A,G; 409R; 411K,T; 412R;
15 417E,R; 421F,Y; 431E.

3. The phytase variant of any of claims 1 or 2, which is derived from an ascomycete phytase.

20 4. The phytase variant of claim 3 which is derived from an *Aspergillus* phytase.

5. The phytase variant of claim 4, wherein the model phytase is a strain of *Aspergillus niger*, *Aspergillus ficum*,
25 *Aspergillus nidulans*, *Aspergillus fumigatus*, *Aspergillus terreus*.

6. The phytase variant of claim 5 wherein the model phytase is *Aspergillus nidulans* DSM 9743; or any of the following
30 strains of *Aspergillus terreus*: CBS 116.46, DSM 9076, CBS 220.95.

7. The phytase variant of claim 6 wherein the model phytase is the *Aspergillus nidulans* phytase sequence shown in Fig. 10; or the *Aspergillus terreus* phytase sequence shown in Fig. 12.

5

8. The phytase variant of claim 3 wherein the model phytase is a strain of *Thermomyces lanuginosus*, *Talaromyces thermophilus*, or *Myceliophthora thermophila*.

10 9. The phytase variant of claim 8 wherein the model phytase is *Thermomyces lanuginosus* CBS 586.94; or any of the following strains of *Talaromyces thermophilus*: ATCC 20186, ATCC 74338; or any of the following strains of *Myceliophthora thermophila*: ATCC 34625, ATCC 74340.

15

10. The phytase variant of claim 9 wherein the model phytase is the *Thermomyces lanuginosus* phytase sequence shown in Fig.14; or the *Talaromyces thermophilus* sequence shown in Fig.13; or the *Myceliophthora thermophila* phytase sequence shown in Fig.7.

20

11. The phytase variant of claim 3 wherein the model phytase is an ascomycete consensus phytase sequence.

12. The phytase variant of any of claims 1 or 2, which is
25 derived from a basidiomycete phytase.

13. The phytase variant of claim 12, wherein the model phytase is a strain of *Paxillus involutus*, *Trametes pubescens*, *Agrocybe pediades*, or *Peniophora lycii*.

30

14. The phytase variant of claim 13 wherein the model phytase is *Trametes pubescens* CBS 100232 or *Paxillus involutus* CBS 100231.
- 5 15. The phytase variant of claim 14 wherein the model phytase is the *Trametes pubescens* phytase sequence of Fig. 4 or either of the *Paxillus involutus* phytase sequences of Figs. 2 and 3.
16. The phytase variant according to any of claims 1 or 2,
 10 which comprises at least one of the following amendments:
 R24C; V27P; H39Q,S; L40N; G42S;
 Q43A,C,D,E,F,G,H,I,K,L,M,N,P,R,S,T,V,W,Y; Y44N; A45D,S; F47Y;
 S49P; A51E,R; V56P; A58D,K; V62I; S69Q; Y75W,F; D78S; S79G;
 K80A; G81A,Q,E; K82T; K83A,I,R,Q; Y84Q,I,V; E90R,A; D115N;
 15 D116S; T118V,L; P119E; F120L; E122A; Q123N,T; L125S,M ; V126H,S;
 N127Q,E; S128A,T; F132I,L; I148V; S151A; S153D,Y; S154Q,D,G;
 I157V; A158D; S159T; G160A,S; K161T,N; K162N; F163W; R170fH;
 Q171N; G173Q,S; S184P,Q; E185S; A186E,P; S187A; T190P,A; P193S;
 G194S,T; T195V,L; V198A,N; E200G,V; D201E; S201d(); E201e(),T;
 20 L201f(); preferably all three deletions; A202S; D203R,K,S;
 D203aV,T; V204Q,E,S,A; T211L,V; S215AP; L220N; D223H; T228N;
 T235Y,L; Y236N; L237F; M238L; S242P; I246V; K251eE,Q; H260A;
 I264R; N265Q,A; Q270Y,A,L,G; S271D,N; K273D; Y275F; H278T;
 A280P; T287A; Q288L,I,F; Y292F; A293V; H302R; P304A; N336S;
 25 G337S,T,Q; I339V; S340P,A; F343A,S,F,I,L; N349P; N360R; T362P;
 F364W; S365V,L,A; S366D,V; A367K; W368K; T369I,L; A373S; S374A;
 R375H; L376M; Q383kE; P404A,G; T411K; R417E; F421Y; A431E.
17. The phytase variant of claim 16, the model phytase of
 30 which is an *Aspergillus* derived phytase, preferably derived from *Aspergillus ficuum* or *Aspergillus niger*.

18. The phytase variant of claim 17, the model phytase of which is a phytase derived from either of *Aspergillus ficuum* (niger) NRRL 3135, *Aspergillus niger* ATCC 9142, or *Aspergillus*
5 *niger* ATCC 74337.

19. The phytase variant of claim 18, the model phytase of which is the *Aspergillus ficuum* phytase sequence of Fig. 11.

10 20. The phytase variant according to any of claims 1 or 2, which phytase variant comprises at least one of the following amendments:

A24C; V27P; H39,S,Q; L40N; G42S; Q43C,D,E,F,H,K,M,P,R,S,W,Y;
Y44N; S45D; F47Y; S49P; E51A,R; L56P; K58D,A; D59G; I62V; S69Q;
15 Y75W,F; S78D; S79G; K80A; S81A,G,Q,E; K82T; K83A,I,Q,R;
Y84Q,V,I; V88K; A90R; F102Y; D115N; D116S; T118V,L; P119E;
F120L; E122A; Q123N,T; L125S,M; V126H,S; N127Q,E; S128A,T;
F132,I,L; S143N; I148V; S151A; S153D,Y; D154Q,S,G; I157V; A158D;
S159T; G160A,S; E161T,N; K162N; F163W; G170fH; ()171N; N173Q,S;
20 T172P; P184Q,S; E185S; S186A,E,P; E187A; T187aS; T190P,A;
G194S,T; V195L,T; K198A,N,V; E200G,V; A201D,E; S201d();
Q201e(),T; L201f(); preferably all three deletions; G202S,A;
D203R,K,S; E203aV,T; V204Q,E,S,A; A205E; L211V; A220L,N; H223D;
T228N; E232T; D233E; V235Y,L,T; V236Y,N; L237F; M238L; C242P,S;
25 T246V; Q251eE,Q; Q256D; H260A; K264R,I; K265Q,A; N267D;
Q270Y,A,L,G; S271D,N; G273D,K; Y275F; Y278T,H; A280P; A287T;
Q288L,I,F; F292Y; T293A,V; R302H; P304A; F332F; N336S;
S337T,G,Q; M338I; V339I; S340P,A; F343A,S,I,L; N349P; E352K;
S360R; K362P; Y364W,F; S365V,L,A; A366D,V,S; S367A,K; W368K;
30 V369I,L; G373S,A; R375H; A376M; K383kQ,E; D404A,G; K411T; I393V;
L412R; K417E,R; W421F,Y; G431E.

21. The phytase variant of claim 20, which is derived from an *Aspergillus* phytase, preferably using a model phytase derived from *Aspergillus fumigatus*.

5

22. The phytase variant of claim 21, the model phytase of which is a phytase derived from either of the following strains of *Aspergillus fumigatus*: ATCC 13073, ATCC 32722, ATCC 58128, ATCC 26906 or ATCC 32239.

10

23. The phytase variant of claim 22, the model phytase of which is the *Aspergillus fumigatus* phytase sequence of Fig. 8.

24. The phytase variant according to any of claims 1 or 2, which phytase variant comprises at least one of the following amendments:

G24C; V27P; H39S,Q; L40N; G42S;

Q43A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y; Y44N; S45D; Y47F;

S49P; E51A,R; V56P; D58K,A; D59G; V62I; S69Q; Y75W,F; S78D;

20 S79G; K80A; S81A,G,Q,E; K82T; A83I,Q,K,R; Y84,Q,I,V; A90R;

D115N; D116S; T118V,L; F119E; P120L; E122A; N123Q,T; M125S;

V126H,S; N127Q,E; S128A,T; Y132F,I,L; K143N; I148V; S151A;

S153D,Y; D154Q,S,G; I157V; A158D; S159T; A160S; E161T,N; K162N;

F163W; G170fH; S170gA; Q171N; H173Q,S; P184Q,S; E185S;

25 G186A,E,P; S187A; G187aS; T190P,A; H193S; G194S,T; T195V,L;

A198N,V; E200G,V; D201E; S201d(); E201e(),T; L201f(); preferably

all three; G202S,A; D203R,K,S; D203aV,T; V204Q,S,A,E; L211V;

A215P; L220N; D223H, T228N; E232T; D233E; V235Y,L,T; Y236N;

L237F; M238L; P242S; E244D; E251e,Q; A256D; H260A; R264I; Q265A;

30 Q270Y,A,L,G; S271D,N; G273D,K; Y275F; Y278T,H; A280P; A287T;

Q288L,I,F; F292Y; A293V; R302H; P304A; N336S; S337T,Q,G; M338I;

I339V; S340P,A; F343A,S,I,L; N349P; A352K; S360R; E362P;
Y364W,F; S365V,L,A; A366D,V,S; S367K,A; W368K; T369I,L; G373S,A;
A374S; R375H; A376M; Q383kE; A404G; K411T; E417R; F421Y; A431E.

5 25. The phytase variant of claim 24, the model phytase of which is an ascomycete consensus phytase.

26. The phytase variant of claim 25, the model phytase of which is the ascomycetes consensus sequence "conphys" of Fig. 9.

10

27. The phytase variant according to any of claims 1 or 2, which phytase variant comprises at least one of the following amendments:

V24C; F27P; ()31Y; F33C; D39H,S,Q; S40L,N; A42S,G;
15 A43C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y; Y44N; T45D,S; Y47F;
Q51E,A,R; K58D,A; K61R; I62V; F75W; S78D; A80K; G81A,Q,E;
R83A,I,Q,K; I84Y,Q,V; V88I; K90R,A; L102Y; D115N; D116S; V118L;
P119E; F120L; L123N,T,Q; S125M; S126H,V; Q127E,N; A128S,T;
T132F,I,L; E143N; V148I; S151A; S152G; S153D,Y; N154D,Q,S,G;
20 D158A; S159T; A160S; T161N; ()170fH; ()170gA; ()171N; H173Q,S;
H172P; S184Q,P; E185S; S186A,E,P; L187A; ()187aS; T190P,A;
D193S; A194S,T; M195T,V,L; N198A,V; G200V; S201D,E ()201eT;
S202A; D203R,K,S; P203aV,T; Q204E,S,A,V; T205E; I211L,V; P215A;
L220N; Q223D,H; A232T; D233E; S235Y,L,T; N236Y; L237F; I238L,M;
25 A242P,S; E244D; I246V; ()251eE,Q; N256D; P260A,H; A264R,I;
Q265A; E267D; G270Y,A,L; L332F; D271N; D273K; F275Y; T278H;
Y280A,P; Y283P; V287A,T; Q288L,I,F; Y292F; I293A,V; E302R,H;
P304A; L332F; N336S; Q337T,S,G; M338I; I339V; A340P;
S343A,F,I,L; F348Y; N349P; S352K; P360R; R362P; W364F;
30 V365L,A,S; T366D,V,S; S367K,A; R368K; L369I; T370V; S373A;

A374S; R375H; S383kQ,E; T387P; A396R; G404A; L409R; T411K;
L412R; E417R; Y421F.

28. The phytase variant of claim 27, the model phytase of
5 which is a phytase derived from *Agrocybe pediades*,

29. The phytase variant of claim 27, the model phytase of
which is a phytase derived from *Agrocybe pediades* CBS 900.96.

10 30. The phytase variant of claim 29, the model phytase of
which is the *Agrocybe pediades* phytase sequence of Fig. 5.

31. The phytase variant according to any of claims 1-2, which
phytase variant comprises at least one of the following
15 amendments:

F24C; V27P; L31Y; I33C; S39H,Q; N40L; G42S;
P43A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y; Y44N; D45S; F47Y;
E51A,R; E58D,K,A; T61R; V62I; W75F; S78D; A80K; R81Q,E,G,A;
S82T; R83A,I,Q,K; Q84Y,V,I; V88I; K90R,A; A115N; D116S; L118V;
20 P119E; F120L; N123T,Q; S125M; H126S,V; Q127E,N; T128A,S;
M132F,I,L; G143N; V148I; A151S; D153Y; Q154D,S,G; D158A; S159T;
S160A; T161N; ()170fH; ()170gA; S171NG172P; E173Q,S; Q184S,P;
E185S; E186A,P; G187A; ()187aS; T190P,A; N193S; N194S,T;
M195T,V,L; N198A,V; V200G; D201E; ()201eT; G202S,A; D203R,K,S;
25 ()203aV,T; E204Q,S,A,V; S205E; V211L; N215A,P; L220N; A223D,H;
S232T; D233E; L235Y,T; T236Y,N; L237F; M238L; P242S; L246V;
()251eE,Q; A260H; V264R,I; S265Q,A; E267D; Y270A,L,G; D271N;
D273K; F275Y; G278T,H; P280A; A283P; T287A; Q288L,I,F; Y292F;
V293A; G302R,H; A304P; N336S; T337Q,S,G; M338I; V339I; P340A;
30 A343S,F,I,L; F348Y; N349P; A352K; E360R; R362P; W364F;

V365L,A,S; D366V,S; S367K,A; L369I; S373A; G374A,S; ()383kQ,E;
E387P; A396R; G404A; V409R; E411K,T; L412R; E417R; Y421F; A431E.

32. The phytase variant of claim 31, the model phytase of
5 which is a phytase derived from *Peniophora lycii*.

33. The phytase variant of claim 32, the model phytase of
which is a phytase derived from *Peniophora lycii* CBS 686.96.

10 34. The phytase variant of claim 33, the model phytase of
which is the *Peniophora lycii* phytase sequence of Fig. 6.

35. A phytase polypeptide which comprises a phytase variant
according to any of the previous claims.

15

36. A DNA construct comprising a DNA sequence encoding a
phytase variant according to any one of claims 1-34.

37. A recombinant expression vector which comprises a DNA
20 construct according to claim 36.

38. A host cell which is transformed with a DNA construct
according to claim 36 or a vector according to claim 37.

25 39. A process for preparing a phytase variant, the process
comprising culturing the host cell according to claim 38 under
conditions permitting the production of the phytase variant, and
recovering the phytase from the culture broth.

30 40. A feed or food comprising at least one phytase variant of
any of claims 1-34.

41. A process for preparing a feed or food according to claim 40, wherein the at least one phytase variant is added to the food or feed components.

5

42. A composition comprising at least one phytase variant of any of claims 1-34.

43. The composition according to claim 42 suitable for use in 10 food or feed preparations.

44. The composition according to any of claims 42-43 which is an animal feed additive.

15 45. A process for reducing phytate levels in animal manure comprising feeding an animal with an effective amount of the feed according to claim 40 or obtainable according to claim 41.

46. Use of the phytase variant of any of claims 1-34; or the 20 composition of any of claims 42-43 for liberating phosphorous from a phytase substrate.

47. A transgenic plant or plant part which is capable of expressing a phytase variant according to any one of claims 1- 25 34.

1/51

Peniophora numbers		1							37
Alignment numbers		1							50
5	P_involtus_A1ML	FGFVALACLL	SLSEVLATSV	P.....KNT	APTFPIESE			
	P_involtus_A2MH	LGFVTLACLI	HLSEVFAASV	P.....RNI	APKFSIESE			
	T_pubescens	MAFSILASLL	FVCYAYARAV	PRAHIPLRDT	SACLDVTRDV			
	A_pediades	MSLFIGGCLL	VFLQASAYGG	VVQATFVQPFFFQI			
	P_lyciiMV	SSAFAPSILL	SLMSSLALST	QFSF.....V	AAQLPIBAQN			
10	A_fumigatusMVTL	TFLLSAAYLL	.SGRVSAAPS	SAGSKSCDTV	DLGYQCSEAT			
	consphyAMGVF	VVLLSIATLF	GSTSGTALGP	RGNSHSCDTV	DGGYQCEPEI			
	A_nidulansMAFF	TVALSLYLL	..SRVSAQAP	VVQNHSCNTA	DGGYQCFENV			
	A_ficuum_NRRL3135MGVS	AVLLPLYLLS	GVTSGLAVPA	SRNQSSCDTV	DQGYQCFSET			
15	A_terreusMGFL	ATVLSVALLF	RSTSGTFLGP	RGKHSDCNSV	DHGYQCFPEL			
	T_thermoMSLL	LLVLSGGLVA	LYVS...RNP	HVDSHSCNTV	EGGYQCRPEI			
	T_lanuginosa		MAGIGLGSFL	VLLLQFSALL	TASPAIPFFW	RKKHPNVD..I		
	M_thermophilaMTGL	GVMVMVGFL	AIASL.....	QSESRPCDTP	DLGFQCGTAI			
		38							83
		51							100
20	P_involtus_A1		QRNWSPYSPY	FPLAEYKA..	..PPAGCQIN	QVNIIQRHGA	RFPTSGATTR		
	P_involtus_A2		QRNWSPYSPY	FPLAEYKA..	..PPAGCEIN	QVNIIQRHGA	RFPTSGAATR		
	T_pubescens		QQSWSMYSPY	FPAATYVA..	..PPASCQIN	QVHIIQRHGA	RFPTSGAAKR		
	A_pediades		QDSWAAATPY	YPVQAYTP..	..PPKDCKIT	QVNIIQRHGA	RFPTSGAGTR		
	P_lycii		TSNWGPYDPF	FPVEPYAA..	..PPEGCTVT	QVNLIQRHGA	RWPTSGARSR		
25	A_fumigatus		SHLWGQYSPF	FSLEDELSVS	SKLPKDCRIT	LVQVLSRHGA	RYPTSSSKSKK		
	consphyA		SHLWGQYSPY	FSLEDESAIS	PDVPDDCRVT	FVQVLSRHGA	RYPTSSSKSKA		
	A_nidulans		SHVWGQYSPY	FSIEQESAIS	EDVPHGCEVT	FVQVLSRHGA	RYPTESKSKA		
	A_ficuum_NRRL3135		SHLWGQYAPF	FSLANESVIS	PEVPAGCRVT	FAQVLSRHGA	RYPTDSKGGK		
30	A_terreus		SHKWGLYAPY	FSLQDESPFP	LDVPEDCHIT	FVQVLSRHGA	RSPTSHSKTA		
	T_thermo		SHSWGQYSPF	FSLADQSEIS	PDVPONCKIT	FVQVLSRHGA	RYPTSSKTEL		
	T_lanuginosa		ARHWGQYSPF	FSLAEVSEIS	PAVPKGCRVE	FVQVLSRHGA	RYPTAHKSEV		
	M_thermophila		SHFWGQYSPY	FSVP..SELD	ASIPDDCEVT	FAQVLSRHGA	RAPTLKRAAS		
		84							133
35	P_involtus_A1		IKAGLTKLQG	VQNFTDAKEN	FIKSFKYDLG	NSDLVPFGAA	QSF DAGQBAF		
	P_involtus_A2		IKAGLSKLQS	VQNFTDPKFD	FIKSFTYDLG	TSDLVPFGAA	QSF DAGLEV		
	T_pubescens		IQTAVAKLKA	ASNYTDPLLA	FVTNYTYSLG	QDSLVELGAT	QSSEAGQBAF		
	A_pediades		IQAAYVKKLQS	AKTYTDPRLD	FLTNYTYTLG	HDDLVPFGAL	QSSQAGEETF		
	P_lycii		QVAAYAKIQM	ARPFTDPKYE	FLNDFVYKFG	VADLLPFGAN	QSHQTGTDMY		
40	A_fumigatus		YKKLVTAIQA	NATDFKKGKFA	FLKTYNYTLG	ADDLTPFGEQ	QLVNSGIKFY		
	consphyA		YSALIEAIQK	NATAFKGKYA	FLKTYNYTLG	ADDLTPFGEN	QMVNSGIKFY		
	A_nidulans		YSGLIEAIQK	NATSFWDGQYA	FLESYNYTLG	ADDLTIFGEN	QMVDSGAKFY		
	A_ficuum_NRRL3135		YSALIEEIQQ	NATTFDGGKYA	FLKTYNYSLG	ADDLTPFGEQ	ELVNSGIKFY		
45	A_terreus		YAATIAAIQK	SATAFPGKYA	FLQSYNYSLD	SEELTPFGRN	QLRDLGAQFY		
	T_thermo		YSQLISRIQK	TATAYKGYA	FLKDYRYQLG	ANDLTPFGEN	QMIQLGIKFY		
	T_lanuginosa		YALLQRIQD	TATEFKGDFA	FLRDYAYHLG	ADNLTRFGEE	QMMESGRQFY		
	M_thermophila		YVDLIDRIHH	GAISYGPGEY	FLRTYDYTLG	ADELTRTGQQ	QMVNSGIKFY		
		134							176
50		151.							200
	P_involtus_A1		ARYSKLVSKN	NLPFIRADGS	DRVVDSATNW	TAGFASA...SHNTVQ		
	P_involtus_A2		ARYSKLVSSD	NLPFIRSDGS	DRVVDATATNW	TAGFASA...SRNAIQ		

2/51

	T_pubescens	TRYSSLVSAD	ELPFVRASGS	DRVVATANNW	TAGFALA...	SSNSIT
	A_pediades	QRYSFVLVSKE	NLPFVRASSS	NRVVDSATNW	TEGFSA...	SHHVLN
	P_lycii	TRYSTLFEGG	DVPFVRAAGD	QRVVDSTTNW	TAGFGDA...	SGETVL
5	A_fumigatus	QRYKAL.ARS	VVPFIRASGS	DRVIASGEKF	IEGFQOAKLA	DPGA.TNRAA
	consphyA	RRYKAL.ARK	IVPFIRASGS	DRVIASAEKF	IEGFQSAKLA	DPGSQPHQAS
	A_nidulans	RRYKNL.ARK	NTPFIRASGS	DRVVASAEKF	INGFRKAQLH	DHGS..KRAT
	A_ficuum_NRRL3135	QRYESL.TRN	IVPFIRSSGS	SRVIASGKKF	IEGFQSTKLK	DPRAQPGQSS
	A_terreus	ERYNAL.TRH	INPFVRATDA	SRVHESAETF	VEGFQATARQD	DHHANPHQPS
	T_thermo	NHYKSL.ARN	AVPFVRCSSG	DRVIASGRLF	IEGFQSAKVL	DPHSDKHDAP
10	T_lanuginosa	HRYREQ.ARE	IVPFVRAAGS	ARVIASAEFF	NRGFQDAKDR	DPRSNDQAE
	M_thermophila	RRYRAL.ARK	SIPFVRTAGQ	DRVVHSAENF	TQGFHSALLA	DRGSTVRPTL
15		177				217
		201				250
	P_involtus_A1	PKLNLILPQT	G..NDTLEDN	MCPAAGD...	...SDPQVNA	WLAVAFPSIT
	P_involtus_A2	PKLDLILPQT	G..NDTLEDN	MCPAAGE...	...SDPQVDA	WLASAFPSVT
	T_pubescens	PVLSVIIESEA	G..NDTLDDN	MCPAAGD...	...SDPQVNO	WLAQFAPPMT
20	A_pediades	PILFVILSES	L..NDTLDDA	MCPNAGS...	...SDPQTGI	WTSIYGTPIA
	P_lycii	PTLQVVLQEE	G..NCTLCNN	MCPNEVD...	...GD.ESTT	WLGVFAPNIT
	A_fumigatus	PAISVIIPES	ETFNNTLDHG	VCTKFEA...	SQLGDEVAAN	FTALFAPDIR
	consphyA	PVIDVIIPEG	SGYNNNTLDHG	TCTAFED...	SELGDDVEAN	FTALFAPAIR
	A_nidulans	PVVNVIIPEI	DGFNNTLDHS	TCVSFEN...	DERADEIEAN	FTAIMGPPIR
25	A_ficuum_NRRL3135	PKIDVVIIESEA	SSSNNTLDPG	TCTVFED...	SELADTVEAN	FTATFVPSIR
	A_terreus	PRVDVAIPEG	SAYNNTLEHS	LCTAFES...	STVGDDAVAN	FTAVFAPAIA
	T_thermo	PTINVIIIEEG	PSYNNNTLDTG	SCPVFED...	SSGGHDAQEK	FAKQFAPAIL
	T_lanuginosa	PVINVIIIEE	TGSNNTLDGL	TCPAEE...	AP.DPTQPAE	FLQVFGPRVL
	M_thermophila	PYDMVVIPEI	AGANNNTLND	LCTAFEEGPy	STIGDDAQDT	YLSTFAGPIT
30		218				252
		251				300
	P_involtus_A1	ARLNAAAPSV	NLTDTDAFNL	VSLCAFLTVS	KEKK.....S
	P_involtus_A2	AQLNAAAPGA	NLTDAFNL	VSLCPFMTVS	KEQK.....S
35	T_pubescens	ARLNAGAPGA	NLTDTDTYNL	LTLCPFETVA	TERR.....S
	A_pediades	NRLNQAPGA	NITAADVSNL	IPLCAFETIV	KETP.....S
	P_lycii	ARLNAAAPSA	NLSDSDALTL	MDMCPFDTLS	SGNA.....S
	A_fumigatus	ARAEXHLPGV	TLTDEDVVS	MDMCSFDTVA	RTSD..ASQ.LS
	consphyA	ARLEADLPGV	TLTDEDVVYL	MDMCPFETVA	RTSD..ATE.LS
40	A_nidulans	KRLNDLPGI	KLTNENVIYL	MDMCSFDTMA	RTAH..GTE.LS
	A_ficuum_NRRL3135	QRLNDLSGV	TLTDTTEVTYL	MDMCSFDTIS	TSTV..DTK.LS
	A_terreus	QRLNDLPGV	QLSTDDVVNL	MAMCPFETVS	LTDD..AHT.LS
	T_thermo	EKIKDHLPGV	DLAVSDVPYL	MDLCPFETLA	RNHT..DT..LS
	T_lanuginosa	KKITKHMPGV	NLTLEDVPLF	MDLCPFDTVG	SDPVLFPQ.LS
45	M_thermophila	ARVNANLPGA	NLTADTVAL	MDLCPFETVA	SSSSDPATAD	AGGGNGRPLS
		253				300
		301				350
	P_involtus_A1	DFCTLFEGIP	GSFEAFAYGG	DLDKFYGTGY	GQELGPVQGV	GYVNELIARL
50	P_involtus_A2	DFCTLFEGIP	GSFEAFAYAG	DLDKFYGTGY	GQALGPVQGV	GYINELLARL
	T_pubescens	EFCDIYEELQ	AE.DAFAYNA	DLDKFYGTGY	GQPLGPVQGV	GYINELLARL
	A_pediades	PFCNLFT..P	EEFAQFEYFG	DLDKFYGTGY	GQPLGPVQGV	GYINELLARL
	P_lycii	PFCDLFT..A	EEYVSIEYYY	DLDKYYGTGP	GNALGPVQGV	GYVNELIARL
	A_fumigatus	PFCQLFT..H	NEWKKYNYLQ	SLGKYYGYGA	GNPLGPAQGI	GFTNELIARL
55	consphyA	PFCALFT..H	DEWRQYDYLQ	SLGKYYGYGA	GNPLGPAQGV	GFTNELIARL
	A_nidulans	PFCALFT..E	KEWLQYDYLQ	SLSKYYGYGA	GSPLGPAQGI	GFTNELIARL

3/51

5	A_ficuum_NRRL3135	PFCDLFT..H	DEWINYDYLO	SLKKYYGHGA	GNPLGPTQGV	GYANELIARL
	A_terreus	PFCDLFT..A	TEWTQYNYLL	SLDKYYGYGG	GNPLGFPVQGV	GWANELMARL
	T_thermo	PFCALST..Q	EEWQAYDYYQ	SLGKYYGNNG	GNPLGPAQGV	GFVNELIARM
	T_lanuginosa	PFCHLFT..A	DDWMAYDYYY	TLDKYYSHGG	GSAFGPSRGV	GFVNELIARM
	M_thermophila	PFCRLFS..E	SEWRAYDYLO	SVGKWWGYGP	GNPLGPTQGV	GFVNELLARL
		301				
		351				
						349
						400
10	P_involtus_A1	TNS.AVRDNT	QTNRTLDASP	VTFPLNKTFF	ADFSHDNLMV	AVFSAMGLFR
	P_involtus_A2	TNS.AVNDNT	QTNRTLDAAP	DTFPLNKTMY	ADFSHDNLMV	AVFSAMGLFR
	T_pubescens	TAQ.NVSDHT	QTNSTLDSSP	ETFPLNRTLY	ADFSHDNQMV	AIFSAMGLFN
	A_pediades	TEM.PVRDNT	QTNRTLDSSP	LTFPLDRSIY	ADLSHDNQMI	AIFSAMGLFN
	P_lycii	TGQ.AVRDET	QTNRTLDSDP	ATFPLNRTFY	ADFSHDNTMV	PIFAALGLFN
15	A_fumigatus	TRS.PVQDHT	STNSTLVSNP	ATFPLNATMY	VDFSHDNSMV	SIFFALGLYN
	consphyA	TRS.PVQDHT	STNHTLDSNP	ATFPLNATLY	ADFSHDNSMI	SIFFALGLYN
	A_nidulans	TQS.PVQDNT	STNHTLDSNP	ATFPLDRKLY	ADFSHDNSMI	SIFFAMGLYN
	A_ficuum_NRRL3135	THS.PVHDDT	SSNHTLDSSP	ATFPLKSTLY	ADFSHDNGII	SILFALGLYN
	A_terreus	TRA.PVHDHT	CVNNTLDASP	ATFPLNATLY	ADFSHDNSLV	SIFWALGLYN
20	T_thermo	THS.PVQDHT	TVNHTLDSNP	ATFPLNATLY	ADFSHDNTMT	SIFAALGLYN
	T_lanuginosa	TGNLPVKDHT	TVNHTLDDNP	ETFPLDAVLY	ADFSHDNTMT	GIFSAMGLYN
	M_thermophila	A.GVPVRDGT	STNRTLDGDP	RTFPLGRPLY	ADFSHDNDMM	GVLGALGAYD
		350				
		401				383
						450
25	P_involtus_A1	QPAPLSTSV	NPWR.....T	WRTSSLVPFS	GRMVVERLSC
	P_involtus_A2	QSAPLSTSTP	DPNR.....T	WLTSSVVPFS	ARMAVERLSC
	T_pubescens	QSAPLDPTTP	DPAR.....T	FLVKKIVPFS	ARMVVERLDC
	A_pediades	QSSPLDPSFP	NPKR.....T	WVTSRLTPFS	ARMVTERLLC	QRDGTGSGGP
	P_lycii	ATA.LDPLKP	DENR.....L	WVDSKLVFPFS	GHMTVEKLAC
30	A_fumigatus	GTEPLSRTSV	ESAKE..LDG	YSASWVVPFG	ARAYFETMQC
	consphyA	GTAPLSTTSV	ESIEE..TDG	YSASWTVPFG	ARAYVEMMQC
	A_nidulans	GTQPLSMDSV	ESIQE..MDG	YAASWTVPFG	ARAYFELMQC
	A_ficuum_NRRL3135	GTKPLSTTTV	ENITQ..TDG	FSSAWTVPFA	SRLYVEMMQC
	A_terreus	GTAPLSQTSV	ESVSQ..TDG	YAAAWTVPFA	ARAYVEMMQC
35	T_thermo	GTAKLSTTEI	KSIEE..TDG	YSAAWTVPFG	GRAYIEMMQC
	T_lanuginosa	GTKPLSTSKI	QPPTGAAADG	YAASWTVPFA	ARAYVELLRC	ETETSSEEEE
	M_thermophila	GVPPLDKTAR	RDPEE..LGG	YAASWAVPFA	ARIYVEKMRC	SGGGGGGGGG
		384				
		451				425
						500
40	P_involtus_A1FGT	TKVRVLVQDQ	VQPLEFCGGD	RNGLCTLAKF	VESQTFARSD
	P_involtus_A2AGT	TKVRVLVQDQ	VQPLEFCGGD	QDGLCALDKF	VESQAYARSG
	T_pubescensGGA	QSVRLLVNDA	VQPLAFCGAD	TSGVCTLDAF	VESQAYARN
	A_pediades	SRIMRNGNVQ	TFVRILVNDA	LQPLKFCGGD	MDSLCTLEAF	VESQKYARED
	P_lyciiSGK	EAVRVLVNDA	VQPLEFCGG	VDGVCELSAF	VESQTYAREN
45	A_fumigatus	K..S...EKE	PLVRALINDR	VVPLHGCDVD	KLGRCKLNDF	VKGLSWARSG
	consphyA	Q..A...EKE	PLVRVLVNDR	VVPLHGCAVD	KLGRCKRDDF	VEGLSFARSG
	A_nidulans	E.....KKE	PLVRVLVNDR	VVPLHGCAVD	KFGRCTLDW	VEGLNFARSG
	A_ficuum_NRRL3135	Q..A...EQA	PLVRVLVNDR	VVPLHGCPVD	ALGRCTRDSF	VRGLSFARSG
	A_terreus	R..A...EKE	PLVRVLVNDR	VMPLHGCPVD	KLGRCKRDAF	VAGLSFAQAG
50	T_thermo	D..D...SDE	PVVRVLVNDR	VVPLHGCEVD	SLGRCKRDDF	VRGLSFARQG
	T_lanuginosa	E..G...EDE	PFVRVLVNDR	VVPLHGCRVD	RWGRCCRDEW	IKGLTFARQG
	M_thermophila	E..GRQEKDE	EMVRVLVNDR	VMTLKGCGAD	ERGMCTLERF	IESMAFARGN
		426				
		439				

4/51

	501	514
P_involtus_A1	GAGDFEKCFA	TSA.
P_involtus_A2	GAGDFEKCLA	TTV.
T_pubescens	GEGDFEKCFA	T...
5 A_pediades	GQGD FEKCFD
P_lycii	GQGDFAKCGF	VPSE
A_fumigatus	..GNWGE CFS
consphyA	..GNWAE CFA	*...
A_nidulans	..GNWKT CFT	L...
10 A_ficuum_NRRL3135	..GDWAE CFA
A_terreus	..GNWAD CF.
T_thermo	..GNWEG CYA	ASE.
T_lanuginosa	..GHWDR CF.
15 M_thermophila	..GKWDL CFA

5/51

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1522 base pairs .
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Paxillus involutus
- (B) STRAIN: CBS 100231

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 58..1383

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 115..1383

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 58..114

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GGATCCGAAT TCGGCACTCG TACGGTCCCC CGGTCTACCC TCTGCTCGCC TTGGAAG	57
ATG CTC TTC GGT TTC GTC GCC CTC GCC TGT CTC TTG TCC CTC TCC GAG	105
Met Leu Phe Gly Phe Val Ala Leu Ala Cys Leu Leu Ser Leu Ser Glu	
-19 -15 -10 -5	
GTC CTT GCG ACC TCC GTG CCC AAG AAC ACA GCG CCG ACC TTC CCC ATT	153
Val Leu Ala Thr Ser Val Pro Lys Asn Thr Ala Pro Thr Phe Pro Ile	
1 5 10	
CCG GAG AGT GAG CAG CGG AAC TGG TCC CCG TAC TCG CCC TAC TTC CCT	201
Pro Glu Ser Glu Gln Arg Asn Trp Ser Pro Tyr Ser Pro Tyr Phe Pro	
15 20 25	
CTT GCC GAG TAC AAG GCT CCT CCG GCG GGC TGC CAG ATC AAC CAG GTC	249
Leu Ala Glu Tyr Lys Ala Pro Pro Ala Gly Cys Gln Ile Asn Gln Val	
30 35 40 45	
AAC ATC ATC CAA AGA CAT GGT GCC CGG TTC CCG ACC TCT GGC GCG ACC	297
Asn Ile Ile Gln Arg His Gly Ala Arg Phe Pro Thr Ser Gly Ala Thr	
50 55 60	
ACC CGT ATC AAG GCG GGT TTG ACC AAG TTG CAA GGC GTC CAG AAC TTT	345
Thr Arg Ile Lys Ala Gly Leu Thr Lys Leu Gln Gly Val Gln Asn Phe	
65 70 75	
ACC GAC GCC AAA TTC AAC TTC ATC AAG TCG TTC AAG TAC GAT CTC GGT	393
Thr Asp Ala Lys Phe Asn Phe Ile Lys Ser Phe Lys Tyr Asp Leu Gly	
80 85 90	
AAC TCG GAC CTC GTT CCG TTC GGT GCA GCA CAG TCC TTC GAC GCT GGT	441
Asn Ser Asp Leu Val Pro Phe Gly Ala Ala Gln Ser Phe Asp Ala Gly	
95 100 105	

6/51

CAG GAG GCC TTC GCC CGC TAC TCG AAG CTT GTC AGC AAG AAC AAC CTG Gln Glu Ala Phe Ala Arg Tyr Ser Lys Leu Val Ser Lys Asn Asn Leu 110 115 120 125	489
CCG TTC ATT CGT GCC GAT GGA AGT GAT CGT GTT GTG GAT TCT GCT ACA Pro Phe Ile Arg Ala Asp Gly Ser Asp Arg Val Val Asp Ser Ala Thr 130 135 140	537
AAC TGG ACT GCG GGT TTC GCT TCG GCA AGT CAC AAC ACG GTC CAG CCC Asn Trp Thr Ala Gly Phe Ala Ser Ala Ser His Asn Thr Val Gln Pro 145 150 155	585
AAG CTG AAC CTG ATT CTC CCG CAA ACT GGC AAT GAT ACC CTG GAA GAT Lys Leu Asn Leu Ile Leu Pro Gln Thr Gly Asn Asp Thr Leu Glu Asp 160 165 170	633
AAT ATG TGC CCT GCT GCT GGC GAT TCT GAC CCC CAG GTC AAC GCG TGG Asn Met Cys Pro Ala Ala Gly Asp Ser Asp Pro Gln Val Asn Ala Trp 175 180 185	681
TTG GCT GTT GCT TTC CCT TCC ATC ACT GCA CGG CTC AAC GCC GCC GCG Leu Ala Val Ala Phe Pro Ser Ile Thr Ala Arg Leu Asn Ala Ala Ala 190 195 200 205	729
CCC TCT GTC AAC CTC ACC GAC ACG GAC GCG TTC AAC CTC GTC AGT CTC Pro Ser Val Asn Leu Thr Asp Thr Asp Ala Phe Asn Leu Val Ser Leu 210 215 220	777
TGC GCT TTC TTG ACA GTC TCG AAG GAG AAG AAG AGT GAC TTC TGC ACC Cys Ala Phe Leu Thr Val Ser Lys Glu Lys Lys Ser Asp Phe Cys Thr 225 230 235	825
CTG TTC GAG GGC ATC CCT GGC TCT TTC GAG GCG TTC GCC TAT GGT GGC Leu Phe Glu Gly Ile Pro Gly Ser Phe Glu Ala Phe Ala Tyr Gly Gly 240 245 250	873
GAC CTT GAC AAG TTC TAC GGT ACC GGT TAC GGT CAG GAA CTC GGA CCC Asp Leu Asp Lys Phe Tyr Gly Thr Gly Tyr Gly Gln Glu Leu Gly Pro 255 260 265	921
GTT CAA GGC GTC GGC TAC GTC AAC GAG CTC ATC GCC CGC CTC ACC AAC Val Gln Gly Val Gly Tyr Val Asn Glu Leu Ile Ala Arg Leu Thr Asn 270 275 280 285	969
TCC GCC GTC CGC GAC AAC ACC CAG ACG AAC CGC ACA CTC GAC GCC TCG Ser Ala Val Arg Asp Asn Thr Gln Thr Asn Arg Thr Leu Asp Ala Ser 290 295 300	1017
CCC GTA ACC TTC CCG TTG AAC AAG ACG TTC TAC GCC GAT TTC TCC CAC Pro Val Thr Phe Pro Leu Asn Lys Thr Phe Tyr Ala Asp Phe Ser His 305 310 315	1065
GAC AAC CTC ATG GTC GCC GTC TTC TCC GCC ATG GGC CTC TTC CGC CAG Asp Asn Leu Met Val Ala Val Phe Ser Ala Met Gly Leu Phe Arg Gln 320 325 330	1113
CCC GCG CCG CTC AGC ACG TCC GTG CCG AAC CCA TGG CGC ACG TGG CGC Pro Ala Pro Leu Ser Thr Ser Val Pro Asn Pro Trp Arg Thr Trp Arg 335 340 345	1161

7/51

ACG AGC TCC CTC GTC CCC TTC TCC GGA CGC ATG GTC GTG GAA CGC CTC Thr Ser Ser Leu Val Pro Phe Ser Gly Arg Met Val Val Glu Arg Leu 350 355 360 365	1209
AGC TGT TTC GGC ACG ACC AAG GTT CGC GTC CTC GTG CAG GAC CAG GTG Ser Cys Phe Gly Thr Thr Lys Val Arg Val Leu Val Gln Asp Gln Val 370 375 380	1257
CAG CCG CTC GAG TTC TGC GGG GGT GAT AGG AAC GGG CTG TGC ACG CTT Gln Pro Leu Glu Phe Cys Gly Gly Asp Arg Asn Gly Leu Cys Thr Leu 385 390 395	1305
GCT AAG TTT GTG GAG AGC CAG ACG TTT GCG AGG AGT GAT GGT GCG GGG Ala Lys Phe Val Glu Ser Gln Thr Phe Ala Arg Ser Asp Gly Ala Gly 400 405 410	1353
GAC TTT GAG AAG TGC TTC GCG ACC TCG GCG TGAGGATGGA CGAACAAAT Asp Phe Glu Lys Cys Phe Ala Thr Ser Ala 415 420	1403
TAAATTGGGG TATTTTATCG TATAATTATG GTGTGTGTAG AACATGGGCT CGGGGTCGAT	1463
GGTGAAAAGC AAAGGTTTAT CGTCTAAAAA AAAAAAAAAA AAAAAATTCC TGCGGCCGC	1522

8/51

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1642 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Paxillus involutus*
- (B) STRAIN: CBS 100231

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 48..1373

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 105..1373

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 48..104

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

```

GGATCCGAAT TCCAGTCCCC AAGCTAATCC TCTGCTCGCC TTGGAAG ATG CAC CTC      56
                                   Met His Leu
                                   -19

GGC TTC GTC ACC CTC GCT TGT CTC ATA CAC CTC TCC GAG GTC TTC GCG      104
Gly Phe Val Thr Leu Ala Cys Leu Ile His Leu Ser Glu Val Phe Ala
-15                      -10                      -5

GCA TCC GTG CCC CGG AAT ATT GCT CCG AAG TTC TCA ATT CCG GAA AGC      152
Ala Ser Val Pro Arg Asn Ile Ala Pro Lys Phe Ser Ile Pro Glu Ser
 1                      5                      10                      15

GAG CAG CGA AAC TGG TCG CCT TAC TCT CCT TAC TTT CCC CTA GCC GAA      200
Glu Gln Arg Asn Trp Ser Pro Tyr Ser Pro Tyr Phe Pro Leu Ala Glu
      20                      25                      30

TAC AAG GCT CCT CCA GCA GGC TGC GAG ATT AAC CAA GTC AAT ATT ATC      248
Tyr Lys Ala Pro Pro Ala Gly Cys Glu Ile Asn Gln Val Asn Ile Ile
      35                      40                      45

CAA CGG CAT GGC GCA CGG TTC CCA ACC TCG GGT GCG GCC ACT CGC ATC      296
Gln Arg His Gly Ala Arg Phe Pro Thr Ser Gly Ala Ala Thr Arg Ile
      50                      55                      60

AAG GCT GGT TTA AGC AAG CTG CAA TCC GTC CAG AAT TTC ACC GAC CCC      344
Lys Ala Gly Leu Ser Lys Leu Gln Ser Val Gln Asn Phe Thr Asp Pro
      65                      70                      75                      80

AAA TTC GAC TTC ATC AAG TCG TTC ACA TAC GAT CTT GGT ACT TCC GAC      392
Lys Phe Asp Phe Ile Lys Ser Phe Thr Tyr Asp Leu Gly Thr Ser Asp
      85                      90                      95

```


9/51

CTC GTG CCA TTC GGC GCA GCA CAA TCA TTC GAT GCC GGC CTG GAG GTC Leu Val Pro Phe Gly Ala Ala Gln Ser Phe Asp Ala Gly Leu Glu Val 100 105 110	440
TTC GCT CGC TAT TCG AAG CTC GTC AGC TCG GAC AAC CTG CCT TTC ATT Phe Ala Arg Tyr Ser Lys Leu Val Ser Ser Asp Asn Leu Pro Phe Ile 115 120 125	488
CGC TCA GAT GGT AGC GAT CGT GTA GTC GAC ACT GCT ACG AAC TGG ACT Arg Ser Asp Gly Ser Asp Arg Val Val Asp Thr Ala Thr Asn Trp Thr 130 135 140	536
GCA GGT TTT GCT TCC GCG AGC CGC AAC GCG ATC CAA CCC AAG CTC GAC Ala Gly Phe Ala Ser Ala Ser Arg Asn Ala Ile Gln Pro Lys Leu Asp 145 150 155 160	584
TTG ATA CTT CCA CAA ACT GGC AAT GAC ACC CTC GAG GAC AAC ATG TGT Leu Ile Leu Pro Gln Thr Gly Asn Asp Thr Leu Glu Asp Asn Met Cys 165 170 175	632
CCA GCT GCT GGC GAA TCC GAC CCT CAG GTC GAT GCG TGG TTG GCG TCC Pro Ala Ala Gly Glu Ser Asp Pro Gln Val Asp Ala Trp Leu Ala Ser 180 185 190	680
GCC TTC CCA TCT GTC ACC GCG CAG CTC AAC GCT GCA GCG CCT GGT GCC Ala Phe Pro Ser Val Thr Ala Gln Leu Asn Ala Ala Pro Gly Ala 195 200 205	728
AAT CTC ACA GAC GCC GAC GCC TTC AAC CTC GTC AGC CTG TGT CCC TTC Asn Leu Thr Asp Ala Asp Ala Phe Asn Leu Val Ser Leu Cys Pro Phe 210 215 220	776
ATG ACA GTT TCG AAG GAG CAG AAG AGC GAC TTC TGC ACG TTG TTC GAG Met Thr Val Ser Lys Glu Gln Lys Ser Asp Phe Cys Thr Leu Phe Glu 225 230 235 240	824
GGA ATC CCT GGA TCG TTC GAG GCG TTT GCC TAT GCC GGC GAC CTT GAC Gly Ile Pro Gly Ser Phe Glu Ala Phe Ala Tyr Ala Gly Asp Leu Asp 245 250 255	872
AAG TTC TAT GGG ACC GGC TAT GGC CAA GCC CTC GGA CCG GTC CAA GGC Lys Phe Tyr Gly Thr Gly Tyr Gly Gln Ala Leu Gly Pro Val Gln Gly 260 265 270	920
GTC GGC TAC ATC AAC GAG CTC CTT GCA CGC CTG ACC AAC TCC GCA GTG Val Gly Tyr Ile Asn Glu Leu Leu Ala Arg Leu Thr Asn Ser Ala Val 275 280 285	968
AAC GAC AAC ACA CAG ACG AAC CGC ACA CTC GAC GCC GCA CCA GAC ACG Asn Asp Asn Thr Gln Thr Asn Arg Thr Leu Asp Ala Ala Pro Asp Thr 290 295 300	1016
TTT CCG CTC AAC AAG ACC ATG TAC GCC GAT TTC TCA CAC GAC AAC CTC Phe Pro Leu Asn Lys Thr Met Tyr Ala Asp Phe Ser His Asp Asn Leu 305 310 315 320	1064
ATG GTC GCC GTG TTC TCC GCC ATG GGC CTC TTC CGC CAA TCC GCA CCG Met Val Ala Val Phe Ser Ala Met Gly Leu Phe Arg Gln Ser Ala Pro 325 330 335	1112

10/51

CTC AGC ACG TCC ACA CCG GAT CCG AAC CGC ACG TGG CTC ACG AGC TCT Leu Ser Thr Ser Thr Pro Asp Pro Asn Arg Thr Trp Leu Thr Ser Ser 340 345 350	1160
GTC GTT CCG TTC TCC GCG CGC ATG GCC GTG GAA CGC CTC AGC TGT GCT Val Val Pro Phe Ser Ala Arg Met Ala Val Glu Arg Leu Ser Cys Ala 355 360 365	1208
GGT ACC ACG AAG GTG CGC GTC CTG GTG CAG GAC CAG GTC CAG CCA CTC Gly Thr Thr Lys Val Arg Val Leu Val Gln Asp Gln Val Gln Pro Leu 370 375 380	1256
GAG TTC TGC GGC GGC GAC CAG GAT GGG TTG TGC GCG CTA GAC AAG TTC Glu Phe Cys Gly Gly Asp Gln Asp Gly Leu Cys Ala Leu Asp Lys Phe 385 390 395 400	1304
GTC GAG AGC CAG GCG TAT GCA CGG AGT GGT GGC GCA GGT GAC TTT GAG Val Glu Ser Gln Ala Tyr Ala Arg Ser Gly Gly Ala Gly Asp Phe Glu 405 410 415	1352
AAG TGT CTT GCG ACG ACG GTG TGAGATGGGG TAATCTACGG TGAAGCAGCG Lys Cys Leu Ala Thr Thr Val 420	1403
GAGAGCCTCT CAACGAATGC AAAGGATAGG TTCGAGGCTT ACTTCATCAA CCTATATCAT	1463
CATAGGACAA GCCCCCAAT AGCCAGACTC GTCGTTTGAC ATCGTGTATG AAAATAACCC	1523
ACCCACGCAC TCCGCTGCCA CTATTCGCGT GTATCGCATA CTAGGCGTTT TCGCCAGTT	1583
GAACATGAGC CCATTCTGTC CCCAGTGAAA AAAAAAAAAA AAAAAATTCC TGCGGCCGC	1642

11/51

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1536 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trametes pubescens*
- (B) STRAIN: CBS 100232

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 79..1407

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 130..1407

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 79..129

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

```

GGATCCGAAT TCGCCCCCAC ATTGTTCCA TCTTAGCAGC CGTCCGCGCC CAGGTCTTCG      60
ATAACCCCCC GCGTGAAT ATG GCC TTC TCA ATC TTG GCC TCG CTG CTC TTC      111
          Met Ala Phe Ser Ile Leu Ala Ser Leu Leu Phe
          -17      -15                      -10

GTG TGT TAT GCA TAC GCC AGG GCT GTG CCC CGT GCA CAT ATC CCG CTC      159
Val Cys Tyr Ala Tyr Ala Arg Ala Val Pro Arg Ala His Ile Pro Leu
   -5                      1                      5                      10

CGC GAC ACC TCC GCG TGT CTA GAT GTA ACA CGC GAT GTG CAG CAG AGC      207
Arg Asp Thr Ser Ala Cys Leu Asp Val Thr Arg Asp Val Gln Gln Ser
          15                      20                      25

TGG TCC ATG TAC TCT CCC TAT TTC CCG GCA GCA ACT TAT GTG GCT CCG      255
Trp Ser Met Tyr Ser Pro Tyr Phe Pro Ala Ala Thr Tyr Val Ala Pro
          30                      35                      40

CCC GCG AGT TGC CAG ATC AAT CAG GTC CAC ATC ATC CAA CGT CAT GGT      303
Pro Ala Ser Cys Gln Ile Asn Gln Val His Ile Ile Gln Arg His Gly
          45                      50                      55

GCA CGC TTT CCC ACG TCT GGC GCA GCA AAG CGC ATC CAG ACA GCA GTA      351
Ala Arg Phe Pro Thr Ser Gly Ala Ala Lys Arg Ile Gln Thr Ala Val
          60                      65                      70

GCG AAG CTG AAG GCC GCG TCC AAC TAC ACC GAT CCC CTG CTC GCG TTC      399
Ala Lys Leu Lys Ala Ala Ser Asn Tyr Thr Asp Pro Leu Leu Ala Phe
          75                      80                      85                      90

GTT ACG AAC TAC ACC TAC AGC TTA GGT CAG GAC AGC CTC GTT GAA CTC      447
Val Thr Asn Tyr Thr Tyr Ser Leu Gly Gln Asp Ser Leu Val Glu Leu
          95                      100                      105

```

12/51

GGT GCG ACT CAG TCC TCC GAA GCG GGC CAG GAG GCA TTC ACG CGG TAC Gly Ala Thr Gln Ser Ser Glu Ala Gly Gln Glu Ala Phe Thr Arg Tyr 110 115 120	495
TCA TCC CTC GTG AGC GCG GAC GAG CTT CCC TTC GTT CGG GCG TCG GGC Ser Ser Leu Val Ser Ala Asp Glu Leu Pro Phe Val Arg Ala Ser Gly 125 130 135	543
TCA GAT CGC GTC GTT GCG ACT GCC AAC AAC TGG ACT GCA GGT TTC GCG Ser Asp Arg Val Val Ala Thr Ala Asn Asn Trp Thr Ala Gly Phe Ala 140 145 150	591
CTT GCG AGC TCA AAC AGC ATC ACG CCC GTG CTC TCA GTC ATC ATT TCC Leu Ala Ser Ser Asn Ser Ile Thr Pro Val Leu Ser Val Ile Ile Ser 155 160 165 170	639
GAA GCG GGC AAT GAC ACC CTC GAC GAC AAC ATG TGC CCC GCT GCA GGC Glu Ala Gly Asn Asp Thr Leu Asp Asp Asn Met Cys Pro Ala Ala Gly 175 180 185	687
GAT TCG GAT CCC CAG GTC AAT CAA TGG CTC GCG CAG TTC GCA CCG CCG Asp Ser Asp Pro Gln Val Asn Gln Trp Leu Ala Gln Phe Ala Pro Pro 190 195 200	735
ATG ACT GCT CGC CTC AAC GCA GGC GCG CCC GGC GCG AAC CTC ACG GAC Met Thr Ala Arg Leu Asn Ala Gly Ala Pro Gly Ala Asn Leu Thr Asp 205 210 215	783
ACG GAC ACC TAC AAC CTG CTC ACG CTA TGC CCG TTC GAG ACT GTA GCC Thr Asp Thr Tyr Asn Leu Leu Thr Leu Cys Pro Phe Glu Thr Val Ala 220 225 230	831
ACC GAG CGG CGT AGT GAA TTC TGC GAC ATC TAC GAG GAG CTG CAG GCG Thr Glu Arg Arg Ser Glu Phe Cys Asp Ile Tyr Glu Glu Leu Gln Ala 235 240 245 250	879
GAA GAC GCC TTC GCG TAC AAT GCC GAT CTC GAC AAG TTC TAC GGC ACT Glu Asp Ala Phe Ala Tyr Asn Ala Asp Leu Asp Lys Phe Tyr Gly Thr 255 260 265	927
GGA TAC GGC CAG CCC CTC GGA CCC GTG CAA GGC GTC GGG TAC ATC AAC Gly Tyr Gly Gln Pro Leu Gly Pro Val Gln Gly Val Gly Tyr Ile Asn 270 275 280	975
GAG CTC ATC GCG CGC CTC ACC GCG CAG AAC GTG TCC GAC CAC ACG CAG Glu Leu Ile Ala Arg Leu Thr Ala Gln Asn Val Ser Asp His Thr Gln 285 290 295	1023
ACG AAC AGC ACA CTC GAC TCC TCG CCC GAG ACG TTC CCG CTC AAC CGC Thr Asn Ser Thr Leu Asp Ser Ser Pro Glu Thr Phe Pro Leu Asn Arg 300 305 310	1071
ACG CTC TAC GCG GAC TTC TCG CAC GAC AAC CAG ATG GTC GCG ATC TTC Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Gln Met Val Ala Ile Phe 315 320 325 330	1119
TCG GCC ATG GGT CTC TTC AAC CAG TCC GCG CCG CTC GAC CCG ACG ACG Ser Ala Met Gly Leu Phe Asn Gln Ser Ala Pro Leu Asp Pro Thr Thr 335 340 345	1167

13/51

CCC GAC CCC GCG CGC ACG TTC CTC GTC AAG AAG ATC GTG CCG TTC TCC Pro Asp Pro Ala Arg Thr Phe Leu Val Lys Lys Ile Val Pro Phe Ser 350 355 360	1215
GCG CGC ATG GTC GTC GAG CGC CTC GAC TGC GGC GGT GCG CAG AGC GTG Ala Arg Met Val Val Glu Arg Leu Asp Cys Gly Gly Ala Gln Ser Val 365 370 375	1263
CGC CTG CTC GTG AAC GAC GCA GTG CAG CCG CTG GCG TTC TGC GGG GCG Arg Leu Leu Val Asn Asp Ala Val Gln Pro Leu Ala Phe Cys Gly Ala 380 385 390	1311
GAC ACG AGC GGG GTG TGC ACG CTG GAC GCG TTT GTC GAG AGC CAG GCG Asp Thr Ser Gly Val Cys Thr Leu Asp Ala Phe Val Glu Ser Gln Ala 395 400 405 410	1359
TAC GCG CGG AAC GAT GGC GAG GGC GAC TTC GAG AAG TGC TTC GCG ACA Tyr Ala Arg Asn Asp Gly Glu Gly Asp Phe Glu Lys Cys Phe Ala Thr 415 420 425	1407
TAGTTCCAGG TGTAGATACC CGGGGAAGAT GTACTCTCTA GACACCTCGC ATGTACTTAT	1467
CGATTAGAAA GAGACCCTGG CTGCTCTGCC CTCAAAAAAA AAAAAAAAAA AAAAAATTCC	1527
TGCGGCCCGC	1536

14/51

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1501 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Agroclybe pediades*
- (B) STRAIN: CBS 900.96

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 17..1375

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 17..94

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 95..1375

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GGATCCGAAT TCACTT ATG TCC CTC TTC ATC GGC GGC TGT TTG CTC GTG	49
Met Ser Leu Phe Ile Gly Gly Cys Leu Leu Val	
-26 -25 -20	
TTT TTA CAG GCG AGC GCA TAC GGC GGC GTC GTG CAG GCC ACA TTC GTG	97
Phe Leu Gln Ala Ser Ala Tyr Gly Gly Val Gln Ala Thr Phe Val	
-15 -10 -5 1	
CAG CCG TTT TTC CCT CCA CAG ATT CAG GAC TCT TGG GCA GCT TAT ACA	145
Gln Pro Phe Phe Pro Pro Gln Ile Gln Asp Ser Trp Ala Ala Tyr Thr	
5 10 15	
CCA TAT TAT CCT GTT CAG GCG TAC ACG CCT CCC CCG AAG GAT TGC AAG	193
Pro Tyr Tyr Pro Val Gln Ala Tyr Thr Pro Pro Pro Lys Asp Cys Lys	
20 25 30	
ATC ACA CAA GTT AAC ATT ATT CAA CGA CAT GGT GCC CGC TTT CCG ACA	241
Ile Thr Gln Val Asn Ile Ile Gln Arg His Gly Ala Arg Phe Pro Thr	
35 40 45	
TCG GGG GCA GGC ACA AGG ATC CAA GCA GCT GTG AAG AAG CTT CAA TCA	289
Ser Gly Ala Gly Thr Arg Ile Gln Ala Ala Val Lys Lys Leu Gln Ser	
50 55 60 65	
GCT AAA ACC TAT ACG GAT CCT CGT CTC GAC TTT CTG ACC AAC TAT ACC	337
Ala Lys Thr Tyr Thr Asp Pro Arg Leu Asp Phe Leu Thr Asn Tyr Thr	
70 75 80	
TAT ACC CTT GGT CAC GAC GAT CTC GTA CCG TTT GGA GCG CTT CAA TCA	385
Tyr Thr Leu Gly His Asp Asp Leu Val Pro Phe Gly Ala Leu Gln Ser	
85 90 95	

15/51

TCA CAA GCT GGA GAG GAA ACG TTT CAA CGA TAC TCG TTT CTG GTG TCC Ser Gln Ala Gly Glu Glu Thr Phe Gln Arg Tyr Ser Phe Leu Val Ser 100 105 110	433
AAA GAG AAC TTA CCT TTT GTA AGA GCT TCG AGT TCC AAT CGA GTC GTC Lys Glu Asn Leu Pro Phe Val Arg Ala Ser Ser Ser Asn Arg Val Val 115 120 125	481
GAC TCA GCT ACC AAC TGG ACG GAA GGT TTT TCT GCG GCC AGT CAC CAC Asp Ser Ala Thr Asn Trp Thr Glu Gly Phe Ser Ala Ala Ser His His 130 135 140 145	529
GTC TTG AAT CCC ATT CTC TTT GTA ATC CTC TCA GAA AGT CTC AAT GAC Val Leu Asn Pro Ile Leu Phe Val Ile Leu Ser Glu Ser Leu Asn Asp 150 155 160	577
ACG CTT GAC GAT GCC ATG TGC CCT AAC GCG GGC TCC TCC GAC CCG CAG Thr Leu Asp Asp Ala Met Cys Pro Asn Ala Gly Ser Ser Asp Pro Gln 165 170 175	625
ACT GGT ATC TGG ACC TCG ATA TAC GGG ACG CCT ATT GCC AAC CGA CTA Thr Gly Ile Trp Thr Ser Ile Tyr Gly Thr Pro Ile Ala Asn Arg Leu 180 185 190	673
AAT CAG CAG GCT CCG GGT GCA AAT ATT ACA GCT GCC GAT GTG TCG AAC Asn Gln Gln Ala Pro Gly Ala Asn Ile Thr Ala Ala Asp Val Ser Asn 195 200 205	721
CTT ATA CCG CTT TGC GCA TTC GAG ACG ATA GTA AAG GAG ACG CCA AGT Leu Ile Pro Leu Cys Ala Phe Glu Thr Ile Val Lys Glu Thr Pro Ser 210 215 220 225	769
CCT TTC TGT AAT TTG TTC ACC CCC GAA GAG TTC GCA CAG TTT GAA TAT Pro Phe Cys Asn Leu Phe Thr Pro Glu Glu Phe Ala Gln Phe Glu Tyr 230 235 240	817
TTC GGT GAC CTG GAC AAG TTC TAT GGG ACA GGT TAT GGA CAA CCG TTA Phe Gly Asp Leu Asp Lys Phe Tyr Gly Thr Gly Tyr Gly Gln Pro Leu 245 250 255	865
GGA CCT GTG CAA GGT GTC GGC TAC ATC AAT GAA CTT CTT GCC CGA CTC Gly Pro Val Gln Gly Val Gly Tyr Ile Asn Glu Leu Leu Ala Arg Leu 260 265 270	913
ACA GAA ATG CCA GTT CGA GAT AAC ACC CAG ACG AAC AGG ACA CTC GAC Thr Glu Met Pro Val Arg Asp Asn Thr Gln Thr Asn Arg Thr Leu Asp 275 280 285	961
TCT TCT CCG CTT ACA TTT CCC CTC GAC CGC AGT ATC TAC GCT GAC CTC Ser Ser Pro Leu Thr Phe Pro Leu Asp Arg Ser Ile Tyr Ala Asp Leu 290 295 300 305	1009
TCG CAC GAT AAC CAA ATG ATC GCG ATA TTT TCA GCG ATG GGT CTT TTC Ser His Asp Asn Gln Met Ile Ala Ile Phe Ser Ala Met Gly Leu Phe 310 315 320	1057
AAC CAG AGT TCA CCT TTG GAT CCG TCC TTC CCC AAC CCC AAG CGT ACT Asn Gln Ser Ser Pro Leu Asp Pro Ser Phe Pro Asn Pro Lys Arg Thr 325 330 335	1105

16/51

TGG GTC ACC AGT CGG CTT ACG CCT TTC AGC GCG AGA ATG GTC ACT GAG	1153
Trp Val Thr Ser Arg Leu Thr Pro Phe Ser Ala Arg Met Val Thr Glu	
340 345 350	
CGG TTG CTG TGT CAA AGG GAT GGG ACA GGG AGC GGT GGA CCA TCC AGG	1201
Arg Leu Leu Cys Gln Arg Asp Gly Thr Gly Ser Gly Gly Pro Ser Arg	
355 360 365	
ATC ATG CGG AAT GGA AAT GTG CAG ACG TTT GTG AGG ATT CTT GTC AAC	1249
Ile Met Arg Asn Gly Asn Val Gln Thr Phe Val Arg Ile Leu Val Asn	
370 375 380 385	
GAT GCT TTA CAG CCT TTG AAG TTC TGC GGA GGG GAC ATG GAT AGT TTG	1297
Asp Ala Leu Gln Pro Leu Lys Phe Cys Gly Gly Asp Met Asp Ser Leu	
390 395 400	
TGT ACT CTG GAA GCG TTC GTC GAG AGC CAG AAG TAT GCA CGA GAG GAT	1345
Cys Thr Leu Glu Ala Phe Val Glu Ser Gln Lys Tyr Ala Arg Glu Asp	
405 410 415	
GGT CAA GGC GAT TTT GAA AAA TGT TTT GAT TAAATATTGC AGTATGCTCA	1395
Gly Gln Gly Asp Phe Glu Lys Cys Phe Asp	
420 425	
GTGAGTAGAC TACAGTGCAG GCCCTGTAAC TCTTGATTG TGTTTCTGGA ATTCCTCGGA	1455
GCGTAGTTTG TAGCAAAAAA AAAAAAAAAA AAATTCCTGC GGCCGC	1501

17/51

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1593 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Peniophora lycii*
- (B) STRAIN: CBS 686.96

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION:123..212

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION:213..1439

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:123..1439

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

```

GGATCCGAAT TCCATCTTCT GCTCTGACCT CCATCTCGCT GAGCGGCCGA CGAGAACCTA      60
GGGGCTCTAA GTCCACGTAC TATCGCCGCG CCTGTGAAGG CCCCATACCA GCCCTTATCG      120
AT ATG GTT TCT TCG GCA TTC GCA CCT TCC ATC CTA CTT AGC TTG ATG      167
Met Val Ser Ser Ala Phe Ala Pro Ser Ile Leu Leu Ser Leu Met
-30          -25          -20

TCG AGT CTT GCT TTG AGC ACG CAG TTC AGC TTT GTT GCG GCG CAG CTA      215
Ser Ser Leu Ala Leu Ser Thr Gln Phe Ser Phe Val Ala Ala Gln Leu
-15          -10          -5          1

CCT ATC CCC GCA CAA AAC ACA AGT AAT TGG GGG CCT TAC GAT CCC TTC      263
Pro Ile Pro Ala Gln Asn Thr Ser Asn Trp Gly Pro Tyr Asp Pro Phe
          5          10          15

TTT CCC GTC GAA CCG TAT GCA GCT CCG CCG GAA GGG TGC ACA GTG ACA      311
Phe Pro Val Glu Pro Tyr Ala Ala Pro Pro Glu Gly Cys Thr Val Thr
          20          25          30

CAG GTC AAC CTG ATT CAG AGG CAC GGC GCG CGT TGG CCC ACA TCC GGC      359
Gln Val Asn Leu Ile Gln Arg His Gly Ala Arg Trp Pro Thr Ser Gly
          35          40          45

GCG CGG TCG CGG CAG GTC GCC GCC GTA GCG AAG ATA CAA ATG GCG CGA      407
Ala Arg Ser Arg Gln Val Ala Ala Val Ala Lys Ile Gln Met Ala Arg
          50          55          60          65

CCA TTC ACG GAT CCC AAG TAT GAG TTC CTC AAC GAC TTC GTG TAC AAG      455
Pro Phe Thr Asp Pro Lys Tyr Glu Phe Leu Asn Asp Phe Val Tyr Lys
          70          75          80

```

18/51

TTC GGC GTC GCC GAT CTG CTA CCG TTC GGG GCT AAC CAA TCG CAC CAA Phe Gly Val Ala Asp Leu Leu Pro Phe Gly Ala Asn Gln Ser His Gln 85 90 95	503
ACC GGC ACC GAT ATG TAT ACG CGC TAC AGT ACA CTA TTT GAG GGC GGG Thr Gly Thr Asp Met Tyr Thr Arg Tyr Ser Thr Leu Phe Glu Gly Gly 100 105 110	551
GAT GTA CCC TTT GTG CGC GCG GCT GGT GAC CAA CGC GTC GTT GAC TCC Asp Val Pro Phe Val Arg Ala Ala Gly Asp Gln Arg Val Val Asp Ser 115 120 125	599
TCG ACG AAC TGG ACG GCA GGC TTT GGC GAT GCT TCT GGC GAG ACT GTT Ser Thr Asn Trp Thr Ala Gly Phe Gly Asp Ala Ser Gly Glu Thr Val 130 135 140 145	647
CTC CCG ACG CTC CAG GTT GTG CTT CAA GAA GAG GGG AAC TGC ACG CTC Leu Pro Thr Leu Gln Val Val Leu Gln Glu Gly Asn Cys Thr Leu 150 155 160	695
TGC AAT AAT ATG TGC CCG AAT GAA GTG GAT GGT GAC GAA TCC ACA ACG Cys Asn Asn Met Cys Pro Asn Glu Val Asp Gly Asp Glu Ser Thr Thr 165 170 175	743
TGG CTG GGG GTC TTT GCG CCG AAC ATC ACC GCG CGA TTG AAC GCT GCT Trp Leu Gly Val Phe Ala Pro Asn Ile Thr Ala Arg Leu Asn Ala Ala 180 185 190	791
GCG CCG AGT GCC AAC CTC TCA GAC AGC GAC GCG CTC ACT CTC ATG GAT Ala Pro Ser Ala Asn Leu Ser Asp Ser Asp Ala Leu Thr Leu Met Asp 195 200 205	839
ATG TGC CCG TTC GAC ACT CTC AGC TCC GGG AAC GCC AGC CCC TTC TGT Met Cys Pro Phe Asp Thr Leu Ser Ser Gly Asn Ala Ser Pro Phe Cys 210 215 220 225	887
GAC CTA TTT ACC GCG GAG GAG TAT GTG TCG TAC GAG TAC TAC TAT GAC Asp Leu Phe Thr Ala Glu Glu Tyr Val Ser Tyr Glu Tyr Tyr Tyr Asp 230 235 240	935
CTC GAC AAG TAC TAT GGC ACG GGC CCC GGG AAC GCT CTC GGT CCT GTC Leu Asp Lys Tyr Tyr Gly Thr Gly Pro Gly Asn Ala Leu Gly Pro Val 245 250 255	983
CAG GGC GTC GGA TAC GTC AAT GAG CTG CTT GCA CGC TTG ACC GGC CAA Gln Gly Val Gly Tyr Val Asn Glu Leu Leu Ala Arg Leu Thr Gly Gln 260 265 270	1031
GCC GTT CGA GAC GAG ACG CAG ACG AAC CGC ACG CTC GAC AGC GAC CCT Ala Val Arg Asp Glu Thr Gln Thr Asn Arg Thr Leu Asp Ser Asp Pro 275 280 285	1079
GCA ACA TTC CCG CTG AAC CGT ACG TTC TAC GCC GAC TTC TCG CAT GAT Ala Thr Phe Pro Leu Asn Arg Thr Phe Tyr Ala Asp Phe Ser His Asp 290 295 300 305	1127
AAC ACC ATG GTG CCC ATC TTT GCG GCG CTC GGG CTC TTC AAC GCC ACC Asn Thr Met Val Pro Ile Phe Ala Ala Leu Gly Leu Phe Asn Ala Thr 310 315 320	1175

19/51

GCC CTC GAC CCG CTG AAG CCC GAC GAG AAC AGG TTG TGG GTG GAC TCT Ala Leu Asp Pro Leu Lys Pro Asp Glu Asn Arg Leu Trp Val Asp Ser 325 330 335	1223
AAG CTG GTA CCG TTC TCT GGA CAT ATG ACG GTC GAG AAG CTG GCA TGT Lys Leu Val Pro Phe Ser Gly His Met Thr Val Glu Lys Leu Ala Cys 340 345 350	1271
TCT GGG AAG GAG GCG GTC AGG GTG CTC GTG AAC GAC GCG GTG CAG CCG Ser Gly Lys Glu Ala Val Arg Val Leu Val Asn Asp Ala Val Gln Pro 355 360 365	1319
CTG GAG TTC TGC GGA GGT GTT GAT GGG GTG TGC GAG CTT TCG GCT TTC Leu Glu Phe Cys Gly Gly Val Asp Gly Val Cys Glu Leu Ser Ala Phe 370 375 380 385	1367
GTA GAG AGC CAG ACG TAT GCG CGG GAG AAT GGG CAA GGC GAC TTC GCC Val Glu Ser Gln Thr Tyr Ala Arg Glu Asn Gly Gln Gly Asp Phe Ala 390 395 400	1415
AAG TGC GGC TTT GTT CCG TCG GAA TAGCGGGAGA CCGTCTATGC TACACAGTAA Lys Cys Gly Phe Val Pro Ser Glu 405	1469
TTGTGTACTC TATAGCACTG TAGCTGTACT TACAAGTCGT AGGGTACGAT CGTACTTACG	1529
CTCGTTTATT GATCCTTCCT TTAAAAAAA AAAAAAAAAA AAAAAAAAAA ATTCCTGCGG	1589
CCGC	1593

20/51

gtcgacgagggcacaaccacgcccgtccctcggcggggtccgagagggccggggtccgggtccga 60
 caaggagacggggcgtcccttcggggcgcggtcgcggtgtgggtgttgctgggacgggtga 120
 ggaggggggacggggcggggcgttgatgacggtacgaatgcgaacgggacacagcgccgtgag 180
 cgtgggtgttgcggttctaatcttcttctggtgggtgtgtacgtgtgggtgtgtatgtgt 240
 ttggggggggggaatgttcttggttaattatcttcttctacccttcttcttcttcttcttct 300
 gttcagcaggtataccccgtgtaggtgtacaggtattatgggacgggtgggtgggtgggtact 360
 acttcttagaaggacgggataaggaaaaagggggaaacacgaatatggcgccctgggtgtgtgc 420
 gtcgagcgtggatgcttgacgcccgtctggcacaacattcttcttcttcttcttcttcttctt 480
 gtacttgatagaggtgttctgggggtcagggcgttgcgctgtgtgttcttaccgaatcaccaac 540
 tagtgctactacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 600
 ctccatttgatacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 660
 gaaacccctctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 720
 ttgcatgcgcatcgttttttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 780
 attgaagactacatatgcgcgaagcgttgacattcaacgggggtcctgcgcgcgcgcgcgcgcgc 840
 gcatgtcgtgtattatccacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 900
 cgggttagagacgggacaggtgagacgggtctcgggttaagcaagcaaatctgggaatgcaaggt 960
 tgaaggtgtaatctgcatagcgtggaaatgagaggggtctcgtgggcagccaggaaggtga 1020
 gacgaaatgaggaaagaggacccagaaagctgtgtgttctggaagtgcgcgcgcgcgcgcgcgc 1080
 caggattaaagtacgggtgtccccatgccaaagctgtgtgttctgaaagcgagtagcggagtag 1140
 tgtccattgttctacgagggatcccccaatgtgttagacatgcctgaatcaattttgtccta 1200
 tttttggattttcaactgtttcttctcgaactgtgtgttctgagcgaactatgccgcaaggtaca 1260
 ctacatgttgtacaataatcatatcatcgaacttctcgtaggaggtgtgtgaaataccgacct 1320
 gctctctcttagcaggtgcctaatggcttctcgtgtactcgatcgaaacgggatcagcaagt 1380
 ccatttgctgttgggtgagatgtacgatttacaacacggtggagaggtgagccacagcgga 1440
 taggcttctggaaggatttctggcgtctcggaaagaggggccactcgcgcccaactaacggcg 1500
 ccgatcttgacatggggtctcgcaggggggtttaaagtgcacactacggagtacggattacac 1560
 agtagtgtatgggtggggggcgaggttgggtggccttgtgtgggtgtcaccgggtgtgtgtgt 1620
 tctcgggggtgtcttggggggcgatgggacccactaaacacgggtgtgtgttggcccggc 1680
 caatctacacggcctcatgttctcggagccagtcagggtgaggtggcactactcagtcagg 1740
 tacacacgtcgggttctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1800
 gttggcataggaggtatcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1860
 cccgggtatctcgttctcctgagcgaagcgcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1920
 cgacggtgcaagaaggttgggggaaggaaatcaaggaaaaaaatagggggggtgg 1980
 ggaccaagagagaaagaaaggagaaaaagggtggggggaggggaagagaaaaaaacggga 2040
 ggaatatggcgctcgtcttctcgaactgttccggaaagggggcatctgggtacacatatgcac 2100
 ctcttccgcacgggaggtatataacggggagtgagtcaccaacggatcatgtgtgtgtgtgt 2160
 cccgtctccagacttccaggtcgagagaggaactagacgcgcggtgaagatgactggcctcg 2220
 M T G L G 5

gagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2280
 V M V V M V G F L A I A S L 19

ggggtccgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2340
 Q S E S R P C 26

ggcacacccacagacttgggttctcagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2400
 D T P D L G F Q C G T A I S H F W G Q Y 46

actcgccttacttctcgtgcccctcggagctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2460
 S P Y F S V P S E L D A S I P D D C E V 66

tgacgttttgcctcagtccttctcggcgaaggcgagggcgccgacgctcaaacggggcg 2520
 T F A Q V L S R H G A R A P T L K R A A 86

cgagctacgtcgatctcatcgacaggtatccaccatggcgccatctcctacggggcggggt 2580
 S Y V D L I D R I H H G A I S Y G P G Y 106

acgagttctcagggactgtactacacccctggcgcgagcgtcaccggagcggggc 2640
 E F L R T Y D Y T L G A D E L T R T G Q 126

agcagcagatggtcaactcgggcatcaagtttaccgcccgtaccggtctctcggcgca 2700
 Q Q M V N S G I K F Y R R Y R A L A R K 146

21/51

agtcgatcccttcgtccgcaccgcccggccaggaccgctcgtccactcggccgagaact 2760
S I P F V R T A G Q D R V V H S A E N F 166

tcacccaggggttccactctgccctgctcgcgcgaccgcsstccaccgtccggccaccc 2820
T Q G F H S A L L A D R G S T V R P T L 186

tccctatgacatggtcgtcatcccggaaccgcccggcccaacaacacgctccacaacg 2880
P Y D M V V I P E T A G A N N T L H N D 206

acctctgcaccgcttcgagggaaggcccgtaactcgaccatcgggcgacgacgcccagaca 2940
L C T A F E E G P Y S T I G D D A Q D T 226

cctacctctccaccttcgcccgaaccatcacgcccgggtcaacyccaacctgcccggcg 3000
Y L S T F A G P I T A R V N A N L P G A 246

ccaacctgaccgacgcccacacggtcgcgctgatggacctctgccccttcgagacggtcg 3060
N L T D A D T V A L H D L C P F E T V A 266

cctcctcctcctccgaccccgcaacggcgacgcccggggcggaacggcgcccgctgt 3120
S S S S D P A T A D A G G G N G R P L S 286

cggccttctgcccgtgttcagcgagtcggagtgggcgcgctacgactacctgcagtcgg 3180
P F C R L F S E S E W R A Y D Y L Q S V 306

tgggcaagtgttacgggtacgggcccgggcaaccgctggggccgacgacgggggtcgggt 3240
G K W Y G Y G P G N P L G P T Q G V G F 326

tgcgcaacgagctgctggcgcggtggcggggtccccgtgcgcgacggcaccagcacca 3300
V N E L L A R L A G V P V R D G T S T N 346

accgcacccctcgacggcgacccggcgacccctccccgtcggccggccctctacgcccact 3360
R T L D G D P R T F P L G R P L Y A D F 366

tcagccacgacaacgacatgatggcggtcctcgggcgccctcgggcgctacgacggcgctcc 3420
S H D N D H M G V L G A L G A Y D G V P 386

cggccctcgacaagaccgcccggcgacccgggaagagctcggcggttacgcccagct 3480
P L D K T A R R D P E E L G G Y A A S W 406

gggcccgtcccgttcgcccgcaggatctacgtcgagaagatgcgggtgcagcggcgccggcg 3540
A V P F A A R I Y V E K H R C S G G G G 426

gcggcgggcgggcgggcgaggggcgggcaggagaaaggatgaggagatgggcagggtgctgg 3600
G G G G G E G R Q E K D E E H V R V L V 446

tgaacgacgggtgatgacgtgaagggtgcccgcgcgacgagagggggatgtgtacgc 3660
N D R V H T L K G C G A D E R G M C T L 466

tagaacgggttcacgaaagcatggcggttcgaggggggaacggcaagtgggatctctgct 3720
E R F I E S H A F A R G N G K W D L C F 486

ttgcttgatatgcccacgcccagatagaacagaacttggatgggggttagagtggtga 3780
A 487

ttcgagatgatagttcacagtcttcgggaatcaaaaatcggttagactggcgaaattcaa 3840
gtctggggcctcgggcgctctgcattctccgttccctgtgttacctcttcaatgggtcttt 3900
ttttattttttattttttttaaattttcacacaaaccttttattgtctttttttctctt 3960
ttttcttctctgcacatcggtatgggaattgtcgac 3995

22/51

1 AGATTCAACGACCGAGGAATCCCAACCCTAATTGTCGGTATCATGGTGAC 50
M V T 3

51 TCTGACTTTCCTGCTTTCGGCGGCTATCTGCTTTCGGGctgagctggctt 100
L T F L L S A A Y L L S G 16

101 ggacccaccgctcgagatagggcttggctggatcccgaaacggagTAGA 150
R 17

151 GTGTCTGCGGCACCTAGTTCTGCTGGCTCCAAGTCCTGCGATACGGTAGA 200
V S A A P S S A G S K S C D T V D 34

201 CCTCGGGTACCAAGTGCTCCCTGCGACTTCTCATCTATGGGGCCAGTACT 250
L G Y Q C S P A T S H L W G Q Y S 51

251 CGCCATTCTTTTCGCTCGAGGACGAGCTGTCCGTGTGAGTAAGCTTCCC 300
P F F S L E D E L S V S S K L P 67

301 AAGGATTGCCGGATCACCTTGGTACAGGTGCTATCGGCCATGGAGCGCG 350
K D C R I T L V Q V L S R H G A R 84

351 GTACCCCAACCAGCTCCAAGAGCAAAAAGTATAAGAAGCTTGTGACGGCGA 400
Y P T S S K S K K Y K K L V T A I 101

401 TCCAGGCCAATGCCACCGACTTCAAGGGCAAGTTTGCTTTTGAAGACG 450
Q A N A T D F K G K F A F L K T 117
+

451 TACAACTATACTCTGGGTGCGGATGACCTCACTCCCTTTGGGGAGCAGCA 500
Y N Y T L G A D D L T P F G E Q Q 134
+

501 GCTGGTGAACTCGGGCATCAAGTTCTACCAGAGGTACAAGGCTCTGGCGC 550
L V N S G I K F Y Q R Y K A L A R 151

551 GCAGTGTGGTGCCGTTTATTCGCGCCTCAGGCTCGGACCGGGTTATTGCT 600
S V V P F I R A S G S D R V I A 157

23/51

601 TCGGGAGAGAAAGTTTCATCGAGCGGTTCCAGCAGGGAAGCTGGCTGATCC 650
S G E K F I E G F Q Q A K L A D P 184

651 TGGCGCGACGAACCGCGCGGCTCCGGCGATTAGTGTGATTATTCGGGAGA 700
G A T N R A A P A I S V I I P E S 201

701 GCGAGACGTTCAACAATACGCTGGACCACGGTGTGTGCACGAAGTTTGAG 750
E T F N N T L D H G V C T K F E 217
+

751 GCGAGTCAGCTGGGAGATGAGGTTGCGGCCAATTTCAGTGGCTCTTTGC 800
A S Q L G D E V A A N F T A L F A 234
+

801 ACCCGACATCCGAGCTCGCGCGGAGAAGCATCTTCCTGGCGTGACGCTGA 850
P D I R A R A E K H L P G V T L T 251

851 CAGACGAGGACGTTGTCACTCTAATGGACATGTGTTGTTTGATACGGTA 900
D E D V V S L M D M C S F D T V 257

901 GCGCGCACCCAGCGACCGCAAGTCAGCTGTTCACCGTTCTGTCAACTCTTCAC 950
A R T S D A S Q L S P F C Q L F T 284

951 TCACAATGAGTGGGAAGTACAACCTACCTTCAGTCTCTGGGCAAGTACT 1000
H N E W K K Y N Y L Q S L G K Y Y 301

1001 ACGGCTACGGCGGAGGCAACCCTCTGGGACCGGCTCAGGGGATAGGGTTC 1050
G Y G A G N P L G P A Q G I G F 317

1051 ACCAACCAGCTGATTGCCCGGTTGACTCGTTCGCCAGTGCAGGACCACAC 1100
T N E L I A R L T R S P V Q D H T 334

1101 CAGCACTAAGTCGACTCTAGTCTCCAACCCGGCCACCTTCCCGTTGAACG 1150
S T N S T L V S N P A T F P L N A 351
+

1151 CTACCATGTACGTGACCTTTTCACACGACAACAGCATGGTTTCCATCTTC 1200
T M Y V D F S H D N S M V S I F 357

1201 TTTGCATTGGGCGCTGTACACGGCACTGAACCCCTGTCCCGGACCTCGGT 1250
F A L G L Y N G T E P L S R T S V 384

24/51

1251 GGAAAGCGCCAAGGAATTGGATGGGTATTCTGCATCCTGGGTGGTGCCTT 1300
E S A K E L D G Y S A S W V V P F 401

1301 TCGGCGCGCGAGCCTACTTTCGAGACGATGCAATGCAAGTCGGAAAAGGAG 1350
G A R A Y F E T M Q C K S E K E 417

1351 CCTCTTGTTTCGCGCTTTGATTAATGACCGGGTTGTGCCACTGCGATGGCTG 1400
P L V R A L I N D R V V P L K G C 434

1401 CGATGTGGACAAGCTGGGGCGATGCAAGCTGAATGACTTTGTCAAGGGAT 1450
D V D K L G R C K L N D F V K G L 451

1451 TGAGTTGGGCCAGATCTGGGGGCAACTGGGGAGAGTGCTTTAGTTGAGAT 1500
S W A R S G G N W G E C F S 465

1501 GTCATTGTTATGCTATACTCCAATAGACCGTTGCTTAGCCATTCACTTCA 1550

1551 CTTTGCTCGAACCGCCTGCCG 1571

25/51

	1		50
<i>A. terreus</i> 9A-1	MGFLAIVLS	VaLLfrsTSG	TFLGprg..K hsGCNSVDhg YQCFPELSHK
<i>A. terreus</i> cbs	MGVFVLLS	IaTLfgsTSG	TALGprg..N hsDCTSVDeG YQCFPELSHK
<i>A. niger</i> var. <i>awamori</i>	MGVsaVLLP	LYLLagVTSG	IaVPasr..N qstCOTVDQG YQCFSETSHL
<i>A. niger</i> T213	MGVsaVLLP	LYLLagVTSG	IaVPasr..N qstCOTVDQG YQCFSETSHL
<i>A. niger</i> NRRL3135	MGVsaVLLP	LYLLsqVTSG	IaVPasr..N qstCOTVDQG YQCFSETSHL
<i>A. fumigatus</i> 13073	MVtLcFLLSa	AYLLsqVSAA	PSsA....G SkSCDTVDIG YQCsPATSHL
<i>A. fumigatus</i> 32722	MVtLcFLLSa	AYLLsqVSAA	PSsA....G SkSCDTVDIG YQCsPATSHL
<i>A. fumigatus</i> 58128	MVtLcFLLSa	AYLLsqVSAA	PSsA....G SkSCDTVDIG YQCsPATSHL
<i>A. fumigatus</i> 26906	MVtLcFLLSa	AYLLsqVSAA	PSsA....G SkSCDTVDIG YQCsPATSHL
<i>A. fumigatus</i> 32239	MGaLcFLLSV	mYLLsqVAGA	PSsGcsagsG SkACDTVELG YQCsPGTSHL
<i>A. nidulans</i>	MAFFcVaLSL	yYLLs=VSAQ	APVV....Q NKSNTADGG YQCFPNVSHV
<i>T. thermophilus</i>	MSLLILVLSg	GLValyVSrN	PHV.....D SHSNTVEGG YQCrPEISHs
<i>M. thermophila</i>	MVGfIAIaSLqSE SRPCDTp0LG FQCgTAISHF
Consensus	MGFL-VLLSL	GYLL--VSAG	PFVG-----N SHSCDTVDGG YQCFPEISHL
Conphys	MGVFVLLS	IATLFGSTAG	YALGPRG..N SHSCDTVDGG YQCFPEISHL

	51		100
<i>A. terreus</i> 9A-1	WGlyAPYFSL	QCESPFPIOV	PEDCHITFVQ VLARHGArSP ThSKcKAYAA
<i>A. terreus</i> cbs	WGlyAPYFSL	QCESPFPIOV	PDDCHITFVQ VLARHGArSP TOSKcKAYAA
<i>A. niger</i> var. <i>awamori</i>	WGQYAPFFSL	ANESAIspDV	PAGCRVTFAQ VLSRHGARYP TESKqKkYSA
<i>A. niger</i> T213	WGQYAPFFSL	ANESVISpDV	PAGCRVTFAQ VLSRHGARYP TESKqKkYSA
<i>A. niger</i> NRRL3135	WGQYAPFFSL	ANESVISPEV	PAGCRVTFAQ VLSRHGARYP TOSKqKkYSA
<i>A. fumigatus</i> 13073	WGQYSPFFSL	EDELsvSSKL	PKDCRITLVQ VLSRHGARYP TSSKsKkYKk
<i>A. fumigatus</i> 32722	WGQYSPFFSL	EDELsvSSKL	PKDCRITLVQ VLSRHGARYP TSSKsKkYKk
<i>A. fumigatus</i> 58128	WGQYSPFFSL	EDELsvSSKL	PKDCRITLVQ VLSRHGARYP TSSKsKkYKk
<i>A. fumigatus</i> 26906	WGQYSPFFSL	EDELsvSSKL	PKDCRITLVQ VLSRHGARYP TSSKsKkYKk
<i>A. fumigatus</i> 32239	WGQYSPFFSL	EDELsvSSDL	PKDCRVTFVQ VLSRHGARYP TASKsKkYKk
<i>A. nidulans</i>	WGQYSPYFSI	EQESAISpDV	PHGCEVTFVQ VLSRHGARYP TESKsKAYSG
<i>T. thermophilus</i>	WGQYSPFFSL	ADQSEISpOV	FQNCKITFVQ VLSRHGARYP TSSKcELYSQ
<i>M. thermophila</i>	WGQYSPYFSV	pSEIDaS..I	PDDCEVTFVQ VLSRHGARaP TLKRaasYvD
Consensus	WGQYSPYFSL	EDESAISpDV	PDDCRVTFVQ VLSRHGARYP TSSK-KAYSA
Conphys	WGQYSPYFSL	EDESAISpDV	PDDCRVTFVQ VLSRHGARYP TSSKsKAYSA

	101		150
<i>A. terreus</i> 9A-1	tIAAIQKSAT	aFpGKYAFLQ	SYNYSLOSEE LTPFGrNQLr DlgAQFYeRY
<i>A. terreus</i> cbs	tIAAIQKNAT	aLpGKYAFLK	SYNYSMGSEN LTPFGrNQLq DlgAQFYRRY
<i>A. niger</i> var. <i>awamori</i>	LIEEIQQNVT	tFDGKYAFLK	TYNYSLGADD LTPFGEQELV NSGIKfYQRY
<i>A. niger</i> T213	LIEEIQQNVT	tFDGKYAFLK	TYNYSLGADD LTPFGEQELV NSGIKfYQRY
<i>A. niger</i> NRRL3135	LIEEIQQNAT	tFDGKYAFLK	TYNYSLGADD LTPFGEQELV NSGIKfYQRY
<i>A. fumigatus</i> 13073	LVTAIQaNAT	dFKGKFAFLK	TYNyTLGADD LTPFGEQQLV NSGIKfYQRY
<i>A. fumigatus</i> 32722	LVTAIQaNAT	dFKGKFAFLK	TYNyTLGADD LTPFGEQQLV NSGIKfYQRY
<i>A. fumigatus</i> 58128	LVTAIQaNAT	dFKGKFAFLK	TYNyTLGADD LTPFGEQQLV NSGIKfYQRY
<i>A. fumigatus</i> 26906	LVTAIQaNAT	dFKGKFAFLK	TYNyTLGADD LTAfGEQQLV NSGIKfYQRY
<i>A. fumigatus</i> 32239	LVTAIQaNAT	eFKGKFAFLK	TYNyTLGADD LTPFGEQQMv NSGIKfYQRY
<i>A. nidulans</i>	LIEAIQKNAT	sFWGQYAFLE	SYNYTLGADD LTfFGENQMv OSGaKfYRRY
<i>T. thermophilus</i>	LISrIQKTAT	aYKGyYAFLE	OYrYqLGAND LTfFGENQMI QLGIKfYnHY
<i>M. thermophila</i>	LIDrIHhGAI	sYgPqYEFLE	TYDYTLGADE LTrcGQQQMv NSGIKfYRRY
Consensus	LIEAIQKNAT	-FKGKYAFLK	TYNyTLGADD LTfFGENQMv NSGIKfYRRY
Conphys	LIEAIQKNAT	aFKGKYAFLK	TYNyTLGADD LTPFGENQMv NSGIKfYRRY

26/51

	151		200
A. terreus 9A-1	NALTRhInPF	VRATDASRVh	ESAEKfVEGF QTARqDOHhA nPHQPSPrVd
A. terreus cbs	DTLTrhInPF	VRAADSSRVh	ESAEKfVEGF QNARqGDPhA nPHQPSPrVd
A. niger var. awamori	ESLTrNIIPF	IRSSGSSRVi	ASGEKfIEGF QSTKLkDP-A qPgQSSPkId
A. niger T213	ESLTrNIIPF	IRSSGSSRVi	ASGEKfIEGF QSTKLkDP-A qPgQSSPkId
A. niger NRRL3135	ESLTrNIIPF	IRSSGSSRVi	ASGEKfIEGF QSTKLkDP-A qPgQSSPkId
A. fumigatus 13073	KALARSVVPF	IRASGSORVI	ASGEKfIEGF QqAKLADPGA .TNRAAPAIIs
A. fumigatus 32722	KALARSVVPF	IRASGSORVI	ASGEKfIEGF QqAKLADPGA .TNRAAPAIIs
A. fumigatus 58128	KALARSVVPF	IRASGSORVI	ASGEKfIEGF QqAKLADPGA .TNRAAPAIIs
A. fumigatus 26906	KALARSVVPF	IRASGSORVI	ASGEKfIEGF QqAKLADPGA .TNRAAPAIIs
A. fumigatus 32239	KALAcSVVPF	IRSSGSSRVi	ASGEKfIEGF QqANVADPGA .TNRAAPVIIs
A. nidulans	KNLARKnTPF	IRASGSORVV	ASAEKfINGF RKAQLhDHGS .gQATPVVn
T. thermophilus	KSLARNAVPF	VRCSGSORVI	ASGrLFIZGF QSAKVlOPhS dKHDApPTIn
M. thermophila	BALARKsIPF	VRTAGqORVV	hSAENFTQGF HSALLADRGs tVRPTLPydm

Consensus KALARKIVPF IRASGSORVI ASAEKfIEGF QSAKLADPGS -PHQASPVI-
 Conphys KALARKIVPF IRASGSORVI ASAEKfIEGF QSAKLADPGS QPHQASPVId

	201		250
A. terreus 9A-1	VaIPEGSAYN	NTLEHSICTA	FES...StVG DDAvANFTAV FAPAIaQRLE
A. terreus cbs	VVIPEGTAYN	NTLEHSICTA	FEA...StVG DAAaDNFTAV FAPAIakRLE
A. niger var. awamori	VVISEASSsN	NTLDPGTCTV	FED...SELA DTVEANFTAT FAPSIRQRLE
A. niger T213	VVISEASSsN	NTLDPGTCTV	FED...SELA DTVEANFTAT FAPSIRQRLE
A. niger NRRL3135	VVISEASSsN	NTLDPGTCTV	FED...SELA DTVEANFTAT FVPSIRQRLE
A. fumigatus 13073	VIIPESETFN	NTLDHGVCTk	FEA...SQLG DEVaANFTAL FAPDIRARaE
A. fumigatus 32722	VIIPESETFN	NTLDHGVCTk	FEA...SQLG DEVaANFTAL FAPDIRARaE
A. fumigatus 58128	VIIPESETFN	NTLDHGVCTk	FEA...SQLG DEVaANFTAL FAPDIRARaE
A. fumigatus 26906	VIIPESETFN	NTLDHGVCTk	FEA...SQLG DEVaANFTAL FAPDIRARaE
A. fumigatus 32239	VIIPESETYN	NTLDHGVCTN	FEA...SELG DEVEANFTAL FAPAIRARIE
A. nidulans	VIIPEIDGFN	NTLDHSTCVS	FEN...DErA DEIEANFTAI MGPPIRkRLE
T. thermophilus	VIIeEGPSYN	NTLDtGSCPv	FED...SSqG HDAQEKfakq FAPAIIEKIK
M. thermophila	VVIPETAGaN	NTLHNDICTA	FEEgpyStIG ODAQDTYlST FAGPITARVN

Consensus VIIPEGSGYN NTLDHGTCTA FED---SELG DDAEANFTAT FAPAIRARLE
 Conphys VIIPEGSGYN NTLDHGTCTA FED...SELG DDVEANFTAL FAPAIRARLE

	251		300
A. terreus 9A-1	ADLPGVqLST	DDVVnLMAMC	PFETVSITD. DAHTLSPFCD
A. terreus cbs	ADLPGVqLSA	DDVVnLMAMC	PFETVSITD. DAHTLSPFCD
A. niger var. awamori	NOLSGVTLTD	TEVTyLMDMC	SFDTISSt. vOTKLSPFCD
A. niger T213	NOLSGVTLTD	TEVTyLMDMC	SFDTISSt. vOTKLSPFCD
A. niger NRRL3135	NOLSGVTLTD	TEVTyLMDMC	SFDTISSt. vOTKLSPFCD
A. fumigatus 13073	kHLPGVTLTD	EDVVSLMDMC	SFDTVARTS. DASQLSPFCQ
A. fumigatus 32722	kHLPGVTLTD	EDVVSLMDMC	SFDTVARTS. DASQLSPFCQ
A. fumigatus 58128	kHLPGVTLTD	EDVVSLMDMC	SFDTVARTS. DASQLSPFCQ
A. fumigatus 26906	kHLPGVTLTD	EDVVSLMDMC	SFDTVARTS. DASQLSPFCQ
A. fumigatus 32239	kHLPGVqLTD	DDVVSLMDMC	SFDTVARTS. DASQLSPFCA
A. nidulans	NOLPGIKLTN	ENVIyLMDMC	SFDTMARTA. HGTELSPFCA
T. thermophilus	DHLPGVDLAv	SDVpyLMDLC	PFETLARNh. TOT.LSPFCA
M. thermophila	ANLPGANLTD	ADTVaLMDLC	PFETVAsSSs dpatadaggg NGRpLSPFcr

Consensus ADLPGVTLTD EDVV-LMDMC PFETVARTS- - - - - DATELSPFCA
 Conphys ADLPGVTLTD EDVVyLMDMC PFETVARTS. DATELSPFCA

27/51

	301		350
<i>A. terreus</i> 9A-1	LFTatEWtQY	NYL1SLDKYY	GYGGGNPLGP VQGVGWaNEL MARLTRAPVH
<i>A. terreus</i> cbs	LFTaaEWtQY	NYL1SLDKYY	GYGGGNPLGP VQGVGWaNEL IARLTRSPVH
<i>A. niger</i> var. <i>awamori</i>	LFTHdEWiHY	DYLSLkKYY	GHGAGNPLGP TQGVGYaNEL IARLTHSPVH
<i>A. niger</i> T213	LFTHdEWiHY	DYLSLkKYY	GHGAGNPLGP TQGVGYaNEL IARLTHSPVH
<i>A. niger</i> NRRL3135	LFTHdEWiNY	DYLSLkKYY	GHGAGNPLGP TQGVGYaNEL IARLTHSPVH
<i>A. fumigatus</i> 13073	LFTHnEWkKY	NYLQSLGKYY	GYGAGNPLGP AQQIGFtNEL IARLTRSPVQ
<i>A. fumigatus</i> 32722	LFTHnEWkKY	NYLQSLGKYY	GYGAGNPLGP AQQIGFtNEL IARLTRSPVQ
<i>A. fumigatus</i> 58128	LFTHnEWkKY	NYLQSLGKYY	GYGAGNPLGP AQQIGFtNEL IARLTRSPVQ
<i>A. fumigatus</i> 26906	LFTHnEWkKY	NYLQSLGKYY	GYGAGNPLGP AQQIGFtNEL IARLTRSPVQ
<i>A. fumigatus</i> 32239	LFTHnEWkKY	DYLSLkKYY	GYGAGNPLGP AQQIGFtNEL IARLTRSPVQ
<i>A. nidulans</i>	IFTEkEWlQY	DYLSLkKYY	GYGAGNPLGP AQQIGFtNEL IARLTRSPVQ
<i>T. thermophilus</i>	LSTQeEWqay	DYLSLkKYY	GnGGGNPLGP AQQGVGFvNEL IARLTHSPVQ
<i>M. thermophila</i>	LFSEsEWraY	DYLSVgKWY	GYGPGNPLGP TQGVGFvNEL LARLaqvPVR

Consensus	LFTH-EW-OY	DYLSLkKYY	GYGAGNPLGP	AQQGVGF-NEL	IARLTRSPVQ
Conphys	LFTHDEWRQY	DYLSLkKYY	GYGAGNPLGP	AQQGVGFANEL	IARLTRSPVQ

	351		400
<i>A. terreus</i> 9A-1	DHTCVNNTLD	ASPATFPLNA	TLYADFSDS NLVSIFWALG LYNGTAPLSq
<i>A. terreus</i> cbs	DHTCVNNTLD	ANPATFPLNA	TLYADFSDS NLVSIFWALG LYNGTkPLSq
<i>A. niger</i> var. <i>awamori</i>	OOTSSNHTLD	SNPATFPLNS	TLYADFSDN GIISILFALG LYNGTkPLST
<i>A. niger</i> T213	OOTSSNHTLD	SNPATFPLNS	TLYADFSDN GIISILFALG LYNGTkPLST
<i>A. niger</i> NRRL3135	OOTSSNHTLD	SSPATFPLNS	TLYADFSDN GIISILFALG LYNGTkPLST
<i>A. fumigatus</i> 13073	DHTSTNStLv	SNPATFPLNA	TMVDFSDN SMVSIFFALG LYNGTEPLSf
<i>A. fumigatus</i> 32722	DHTSTNStLv	SNPATFPLNA	TMVDFSDN SMVSIFFALG LYNGTGPLSf
<i>A. fumigatus</i> 58128	DHTSTNStLv	SNPATFPLNA	TMVDFSDN SMVSIFFALG LYNGTEPLSf
<i>A. fumigatus</i> 26906	DHTSTNStLv	SNPATFPLNA	TMVDFSDN SMVSIFFALG LYNGTEPLSf
<i>A. fumigatus</i> 32239	DHTSTNStLD	SDPATFPLNA	TIYDFSDN GMIPIFFAMG LYNGTEPLSq
<i>A. nidulans</i>	ONTSTNHTLD	SNPATFPLDf	KLYADFSDN SMISIFFAMG LYNGTQPLSm
<i>T. thermophilus</i>	DYTTVNHTLD	SNPATFPLNA	TLYADFSDN TMTSIFaALG LYNGTAKLST
<i>M. thermophila</i>	DqTSTNRTLD	GDPfTFPLGr	PLYADFSDN DMMGVLqALG ayDGVFPPLDK

Consensus	DHTSTNHTLD	SNPATFPLNA	TLYADFSDN	SMISIFFALG	LYNGTAPLST
Conphys	DHTSTNHTLD	SNPATFPLNA	TLYADFSDN	SMISIFFALG	LYNGTAPLST

	401		450
<i>A. terreus</i> 9A-1	TSVESVSQTD	GYAAAWTVPF	AARAYVEMMQ C.....RAEKEP
<i>A. terreus</i> cbs	TTVEDITrTD	GYAAAWTVPF	AARAYIEMMQ C.....RAEKQP
<i>A. niger</i> var. <i>awamori</i>	TTVENITQTD	GFSSAWTVPF	ASRLYVEMMQ C.....QAEQEP
<i>A. niger</i> T213	TTVENITQTD	GFSSAWTVPF	ASRLYVEMMQ C.....QAEQEP
<i>A. niger</i> NRRL3135	TTVENITQTD	GFSSAWTVPF	ASRLYVEMMQ C.....QAEQEP
<i>A. fumigatus</i> 13073	TSVESaKElD	GYSASWVVPF	GARAYFetMQ C.....KSEKEP
<i>A. fumigatus</i> 32722	TSVESaKElD	GYSASWVVPF	GARAYFetMQ C.....KSEKEP
<i>A. fumigatus</i> 58128	TSVESaKElD	GYSASWVVPF	GARAYFetMQ C.....KSEKES
<i>A. fumigatus</i> 26906	TSVESaKElD	GYSASWVVPF	GARAYFetMQ C.....KSEKEP
<i>A. fumigatus</i> 32239	TSaESTKESN	GYSASWAVPF	GARAYFetMQ C.....KSEKEP
<i>A. nidulans</i>	DSVESIQEmD	GYAASWTVPF	GARAYFELMQ C.....E.KKEP
<i>T. thermophilus</i>	TEIKSIEETO	GYSAAWTVPF	GGRAYIEMMQ C.....DDSOEP
<i>M. thermophila</i>	TArDpEElG	GYAASWAVPF	AARIYVEKMR Csggggggggg gegrQEKEDe

Consensus	TSVESIEETO	GYSASWTVPF	GARAYVEMMQ	C-----	----QAEKEP
Conphys	TSVESIEETD	GYSASWTVPF	GARAYVEMMQ	C.....QAEKEP

28/51

	451		500
A. terreus 9A-1	LVRVLVNDRV MPLHGCPIDK LGRCK-DAFV AGLSFAQAGG NWADCF----		
A. terreus cbs	LVRVLVNDRV MPLHGCAVDN LGRCK-DDFV EGLSFARAGG NWAECF----		
A. niger var. awamori	LVRVLVNDRV VPLHGCPIDa LGRCT-DSFV rGLSFARSGG DWAECSA----		
A. niger T213	LVRVLVNDRV VPLHGCPIDa LGRCT-DSFV rGLSFARSGG DWAECSA----		
A. niger NRRL3135	LVRVLVNDRV VPLHGCPVDA LGRCT-DSFV rGLSFARSGG DWAECSA----		
A. fumigatus 13073	LVRALINDRV VPLHGCDVDK LGRCKLNDFV KGLSWARSGG NWGECFS----		
A. fumigatus 32722	LVRALINDRV VPLHGCDVDK LGRCKLNDFV KGLSWARSGG NWGECFS----		
A. fumigatus 58128	LVRALINDRV VPLHGCDVDK LGRCKLNDFV KGLSWARSGG NWGECFS----		
A. fumigatus 26906	LVRALINDRV VPLHGCDVDK LGRCKLNDFV KGLSWARSGG NWGECFS----		
A. fumigatus 32239	LVRALINDRV VPLHGCAVDK LGRCKLNDFV KGLSWARSGG NSEQSFS----		
A. nidulans	LVRVLVNDRV VPLHGCAVDK FGRCTLDDWV EGLNFAFARSGG NWKTCFTL--		
T. thermophilus	VVRVLVNDRV VPLHGCEVDS LGRCK-DDFV rGLSFARSGG NWECCYAase		
M. thermophila	MVRVLVNDRV MTLkGCCADE rGMCTLrFI ESMAFARGNG KWDLCFA----		
Consensus	LVRVLVNDRV VPLHGCAVDK LGRCK-DDFV EGLSFARSGG NWAECFA----		
Conphys	LVRVLVNDRV VPLHGCAVDK LGRCK-DDFV EGLSFARSGG NWAECFA		

29/51

CP-1

TATATGAATTCATGGGCGTGTTTCGTCTGCTACTGTCCATTGCCACCTTGTTTCGGTTCCA
1 -----+-----+-----+-----+-----+-----+ 60
ATATACTTAAGTACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGT
CATCCGGTACCGCCTTGGGTCTCTCGTGGTAATTCTCACTCTTGTGACACTGTTGACGGTG
61 -----+-----+-----+-----+-----+-----+ 120
GTAGGCCATGGCGGAACCCAGGAGCACCATTAAAGAGTGAGAACTGTGACAACTGCCAC

CP-2

CP-3

GTTACCAATGTTTCCCAGAAATTTCTCACTTGTGGGGTCAATACTCTCCATACTTCTCTT
121 -----+-----+-----+-----+-----+-----+ 180
CAATGGTTACAAAGGGTCTTTAAAGAGTGAACACCCCAAGTTATGAGAGGTATGAAGAGAA
TGGAAGACGAATCTGCTATTTCTCCAGACGTTCCAGACGACTGTAGAGTTACTTTCGTTT
181 -----+-----+-----+-----+-----+-----+ 240
ACCTTCTGCTTAGACGATAAAGAGGTCTGCAAGGTCTGCTGACATCTCAATGAAAGCAAG

CP-4

CP-5

AAGTTTGTCTAGACACGGTGCTAGATACCCAATTCTTCTAAGTCTAAGGCTTACTCTG
241 -----+-----+-----+-----+-----+-----+ 300
TTCAAAACAGATCTGTGCCACGATCTATGGGTTGAAGAAGATTCAGATTCGGAATGAGAC
CTTTGATTGAAGCTATTCAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTCTTGA
301 -----+-----+-----+-----+-----+-----+ 360
GAAACTAACTTCGATAAGTTTCTTTCGATGACGAAAGTTCCCATTCATGCGAAAGAACT

CP-6

CP-7

AGACTTACAACACTACTTTGGGTGCTGACGACTTGACTCCATTCCGGTGAAAACCAAATGG
361 -----+-----+-----+-----+-----+-----+ 420
TCTGAATGTTGATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTTTGGTTTACC
TTAACTCTGGTATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCAT
421 -----+-----+-----+-----+-----+-----+ 480
AATTGAGACCATAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTTCTAACAAGGTA

CP-8

CP-9

TCATTAGAGCTTCTGGTTCTGACAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTT
481 -----+-----+-----+-----+-----+-----+ 540
AGTAATCTCGAAGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAA

30/51

TCCAATCTGCTAAGTTGGCTGACCCAGGTTCTCAACCACACCAAGCTTCTCCAGTTATTG
541 -----+-----+-----+-----+-----+-----+ 600
AGGTTAGACGATTCAACCGACTGGGTCCAAGAGTTGGTGTGGTTCTGAAGAGGTCAATAAC
CP-10

CP-11
ACGTTATTATTCCAGAAGGATCCTGGTTACAACAACACTTTGGACCACGGTACTTGTACTG
601 -----+-----+-----+-----+-----+-----+ 660
TGCAATAATAAGGTCTTCCTAGGCCAATGTTGTTGTGAAACCTGGTGCCATGAACATGAC

CTTTCGAAGACTCTGAATTGGGTGACGACGTTGAAGCTAACTTCACTGCTTTGTTTCGCTC
661 -----+-----+-----+-----+-----+-----+ 720
GAAAGCTTCTGAGACTTAACCCACTGCTGCAACTTCGATTGAAGTGACGAAACAAGCGAG
CP-12

CAGCTATTAGAGCTAGATTGGPAGCTGACTTGCCAGGTGTTACTTTGACTGACGAAGACG
721 -----+-----+-----+-----+-----+-----+ 780
GTCGATAATCTCGATCTAACCTTCGACTGAACGGTCCACAATGAAACTGACTGCTTCTGC
CP-13
TTGTTTACTTGATGGACATGTGTCCATTGCGAACTGTTGCTAGAACTTCTGACGCTACTG
781 -----+-----+-----+-----+-----+-----+ 840
AACAAATGAACCTACCTGTACACAGGTAAGCTTTGACAACGATCTTGAAGACTGCGATGAC

AATTGTCTCCATTCTGTGCTTTGTTCACTCAGCAGGAATGGAGACAATACGACTACTTGC
841 -----+-----+-----+-----+-----+-----+ 900
TTAACAGAGGTAAGACACGAAACAAGTGAGTGCTGCTTACCTCTGTTATGCTGATGAACG
CP-14

CP-15
AATCTTTGGGTAAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTG
901 -----+-----+-----+-----+-----+-----+ 960
TTAGAAACCCATTGATGATGCCAATGCCACGACCATTTGGGTAACCCAGGTGAGTTCCAC

TTGGTTTCGCTAACGAATTGATTGCTAGATTGACTAGATCTCCAGTTCAAGACCACACTT
961 -----+-----+-----+-----+-----+-----+ 1020
AACCAAAGCGATTGCTTAACTAACGATCTAACTGATCTAGAGGTCAAGTTCTGGTGTGAA
CP-16

CP-17
CTACTAACCACACTTTGGAATCTAACCCAGCTACTTTCCCATTTGAACGCTACTTTGTACG
1021 -----+-----+-----+-----+-----+-----+ 1080
GATGATTGGTGTGAAACCTGAGATTGGGTGCGATGAAAGGGTAACTTGCGATGAAACATGC

31/51

CTGACTTCTCTCAGACAACCTCTATGATTTCATTTTCTTCGCTTTGGGTTTGTACAACG
1081 -----+-----+-----+-----+-----+ 1140
GACTGAAGAGAGTGCTGTTGAGATACTAAAGATAAAAGAAGCGAAACCCAAACATGTTGC

CP-18

CP-19

GTACTGCTCCATTGTCTACTACTTCTGTTGAATCTATTGAAGAACTGACGGTTACTCTG
1141 -----+-----+-----+-----+-----+ 1200
CATGACGAGGTAACAGATGATGAAGACAACCTTAGATAACTTCTTTGACTGCCAATGAGAC

CTTCTTGGACTGTTCCATTGGTGCTAGAGCTTACGTTGAAATGATGCAATGTCAAGCTG
1201 -----+-----+-----+-----+-----+ 1260
GAAGAACCTGACAAGGTAAGCCACGATCTCGAATGCAACTTTACTACGTTACAGTTCGAC

CP-20

CP-21

AAAAGGAACCATTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTG
1261 -----+-----+-----+-----+-----+ 1320
TTTTCTTGGTAAACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACAC

CTGTTGACAAGTTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTGCTA
1321 -----+-----+-----+-----+-----+ 1380
GACAACTGTTCAACCCATCTACATTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGAT

CP-22

GATCTGGTGGTAACTGGGCTGAATGTTTCGCTTAAGAATTCATATA
1381 -----+-----+-----+-----+-----+ 1426
CTAGACCACCATTGACCCGACTTACAAAGCGAATTCTTAAGTATAT

32/51

1	TCTGTAACCGATAGCGGACCGACTAGGCATCGTTGATCCACAATATCTCA	50
51	GACAATGCAACTCAGTCGAATATGAAGGGCTACAGCCAGCATTAAATAC	100
101	GGCCGTCTAGGTCGGGCTCCGGGGATGAGGAGGAGCAGGCTCGTGTTCAT	150
151	TTCCGTCATGGCTTTTTCACGGTCGGCTCTTTCGGCTTTTATTACTTGCTAT	200
	M A F F T V A L S L Y Y L L S	15
201	CGAGGtGagatccctacaaatccctgctccttagtcgaattcggtaccttat	250
	R	16
251	ccgtacagAGTCTCTGCTCAGTCCCCAGTGGTCCAGAATCATTATGCAA	300
	V S A Q A P V V Q N H S C N	30
	+	
301	TACGGCGGACCGGTGGATATCAATGCTTCCCCAATGTCTCTCATGTTGGG	350
	T A D G G Y Q C F F N V S H V W G	47
	+	
351	GTCAGTACTCGCCGTACTTCTCCATCGAGCAGGAGTCAGCTATCTCTGAG	400
	Q Y S P Y F S I E Q E S A I S E	63
401	GACGTGCCCTCATGGCTGTGAGGTTACCTTTGTGCAGGTGCTCTCGCGGCA	450
	D V P H G C E V T F V Q V L S R H	80
451	TGGGGCTAGGTATCCGACAGAGTCGAAGAGTAAGGCGTACTCGGGGTTGA	500
	G A R Y P T E S K S K A Y S G L I	97
501	TTGAAGCAATCCAGAAGAATGCTACCTCTTTTGGGGACAGTATGCTTTT	550
	E A I Q K N A T S F W G Q Y A F	113
	+	
551	CTGGAGAGTTATAACTATACCCTCGGCGCGGATGACTTGACTATCTTCGG	600
	L E S Y N Y T L G A D D L T I F G	130
	+	
601	CGAGAACCAGATGGTTGATTGGGTGCCAAGTTCTACCGACGGTATAAGA	650
	E N Q M V D S G A K F Y R R Y K N	147
651	ATCTCGCCAGGAAAAATACTCCTTTTATCCGTGCATCAGGGTCTGACCGT	700
	L A R K N T P F I R A S G S D R	163

33/51

701 GTCGTTGCGTCTGCGGAGAGTTCATTAAATGGATTTCGCAAGGCTCAGCT 750
 V V A S A E K F I N G F R K A Q L 180
 751 CCACGACCATGGCTCCAAACGTGCTACGCCAGTTGTCAATGTGATTATCC 800
 H D H G S K R A T P V V N V I I P 197
 801 CTGAAATCGATGGGTTTAAACAACCCCTGGACCATAGCACGTGCGTATCT 850
 E I D G F N N T L D H S T C V S 213
 +
 851 TTTGAGAATGATGAGCGGGCGGATGAAATTGAAGCCAATTTACGGCAAT 900
 F E N D E R A D E I E A N F T A I 230
 +
 901 TATGGGACCTCCGATCCGCAACGTCTGGAAAATGACCTCCCTGGCATCA 950
 M G P P I R K R L E N D L P G I K 247
 951 AACTTACAAACGAGAATGTAATATATTGATGGATATGTGCTCTTTCGAG 1000
 L T N E N V I Y L M D M C S F D 263
 1001 ACCATGGCGCGCACCGCCACGGAACCGAGCTGTCTCCATTTGTGCCAT 1050
 T M A R T A H G T E L S P F C A I 280
 1051 CTTCACTGAAAAGGAGTGGCTGCGAGTACGACTACCTTCAATCTCTATCA 1100
 F T E K E W L Q Y D Y L Q S L S K 297
 1101 AGTACTACGGCTACGGTGCCGAAGCCCCCTTGGCCAGCTCAGGGAATT 1150
 Y Y G Y G A G S P L G P A Q G I 313
 1151 GGCTTCACCAACGAGCTGATTGCCCGACTAACGCAATCGCCCGTCCAGGA 1200
 G F T N E L I A R L T Q S P V Q D 330
 1201 CAACACAAGCACCAACCACACTCTAGACTCGAACCCAGCCACATTTCCGC 1250
 N T S T N H T L D S N P A T F P L 347
 +
 1251 TCGACAGGAAGCTCTACGCCGACTTCTCCCAGGACAATAGCATGATATCG 1300
 D R K L Y A D F S H D N S M I S 363
 1301 ATATTCTTCGCCATGGGTCTGTACAACGGCACCCAGCCGCTGTCAATGGA 1350
 I F F A M G L Y N G T Q P L S M D 380
 +

34/51

1351 TTCCGTGGAGTCGATCCAGGAGATGGACGGTTACGGGGCGTCTTGGACTG 1400
S V E S I Q E M D G Y A A S W T V 397

1401 TTCCGTTTGGTGCGAGGGCTTACTTTGAGCTCATGCGTGGGAGAAGAAG 1450
P F G A R A Y F E L M Q C E K K 413

1451 GAGCCGCTTGTGCGGGTATTAGTGAATGATCGCGTGTTCCTCTTCATGG 1500
E P L V R V L V N D R V V P L H G 430

1501 CTGCGCAGTTGACAAGTTTGGACGGTGCACTTTGGACGATTGGGTAGAGG 1550
C A V D K F G R C T L D D W V E G 447

1551 GCTTGAATTTTGCAAGGAGCGGCGGGAAGTGGAGACTGTGTTTACCCTA 1600
L N F A R S G G N W K T C F T L 463

1601 TAAAGGGCGTTTGCTCATTCTATAAGTGTGTGTCAGGTATAGGAAGGTTAG 1650

1651 GCGATTAGCTGTGTTGGCTTTACTCTTATTAGACCAAGAATGATTGTGTTG 1700

1701 TTCTCAAGGCCCTTCTAGCATATCGTCAAGTGGGATAAATCACCTATCCTC 1750

1751 CATGTGTAGGTGAACCCGCTCTTGCAATCAACCTCTTGTTGTTTCAGAGTAG 1800

1801 TTTACCAAACATATCCTCGTGTCTCTCTCTCTGCTCTTCGGTCTCATAT 1850

1851 TAACTGTTCTCTATCTATATCGTCAACAAAATACCACCCAAACACCAA 1900

1901 ATGTCACACTTTCAGCACGAAATTTCTTCG 1911

Fig. 10C

35/51

1 ATGGGCGTCTCTGCTGTTCTACTTCCTTTGTATCTCCTGCTGAGTCACCTCCGGACTG
-23 M G V S A V L L P L Y L L S G V T S G L

61 GCAGTCCCCGCTCGAGAAATCAATCCAGTTGCGATACGGTCGATCAGGGGTATCAATGC
A V P A S R N Q S S C D T V D Q G Y Q C
-1 +1

121 TTCTCCGAGACTTCGCATCTTTGGGGTCAATACGCACCGTTCTTCTCTCTGGCAAACGAA
18 F S E T S H L W G Q Y A P F F S L A N E

181 TCGGTCATCTCCCTGAGGTGCCCGCCGGATGCAGAGTCACTTTGCTCAGGTCTCTCTCC
38 S V I S P E V P A G C R V T F A Q V L S

241 CGTCATGGAGCGCGGTATCCGACCGACTCCAAGGGCAAGAAATACTCCGCTCTCATTGAG
58 R H G A R Y P T D S K G K K Y S A L I E

301 GAGATCCAGCAGAACCGGACCACCTTTGACGGAAAATATGCCTTCTGAGACATACAAC
78 E I Q Q N A T T F D G K Y A F L K T Y N

361 TACAGCTTGGGTGCAGATGACCTGACTCCCTTCGGAGAACAGGAGCTAGTCAACTCCGGC
98 Y S L G A D D L T P F G E Q E L V N S G

421 ATCAAGTTCTACCAGCGGTACGAATCGCTCACAAGGAACATCGTTCCATTGATCCGATCC
118 I K F Y Q R Y E S L T R N I V P F I R S

481 TCTGGCTCCAGCCGCGTGTATCGCTCCGCAAGAAATTCATCGAGGGGTTCCAGAGCACC
138 S G S S R V I A S G K K F I E G F Q S T

541 AAGCTGAAGGATCCTCGTGCCAGCCCGGCCAATCGTCGCCCAAGATCGACGTGGTCAAT
158 K L K D P R A Q P G Q S S P K I D V V I

601 TCCGAGGCCAGCTCATCCAACAACACTCTCGACCCAGGCACCTGCACTGTCTTCAAGAC
178 S E A S S S N N T L D P G T C T V F E D

661 AGCGAATTGGCCGATACCGTCAAGCCAATTTACCGCCACGTTCTGTCCTCCATTCTGT
198 S E L A D T V E A N F T A T F V P S I R

721 CAACGCTCTGGAGAACGACCTGTCCGGTGTGACTCTCACAGACACAGAAGTGACCTACCTC
218 Q R L E N D L S G V T L T D T E V T Y L

781 ATGGACATGTGCTCCTTCGACACCATCTCCACCAGCACCGTCGACACCAAGCTGTCCCCC
238 M D M C S F D T I S T S T V D T K L S P

36/51

841 TTCTGTGACCTGTTACCCCATGACGAATGGATCAACTACGACTACCTCCAGTCCTTGAAA
258 F C D L F T H D E W I N Y D Y L Q S L K

901 AAGTATTACGGCCATGGTGCAGGTAACCCGCTCGGCCCCGACCCAGGGCGTGGGTACGCT
278 K Y Y G H G A G N P L G P T Q G V G Y A

961 AACGAGCTCATCGCCCGTCTGACCCACTCGCCTGTCCACGATGACACCAGTTCCAACCAC
298 N E L I A R L T H S P V H D D T S S N H

1021 ACTTTGGACTCGAGCCCGGCTACCTTTCCGCTCAACTCTACTCTCTACGCGGACTTTTCG
318 T L D S S P A T F P L N S T L Y A D F S

1081 CATGACAACGGCATCATCTCGATTCTCTTTGCTTTAGGTCGTACAACGGCACTAAGCCG
338 H D N G I I S T L F A L G L Y N G T K P

1141 CTATCTACCACGACCGTGGAGAATATCACCCAGACAGATGGATTCTCGTCTGCTTGGACG
358 L S T T T V E N I T Q T D G F S S A W T

1201 GTTCCGTTTGCTTCGCGTTTGTACGTCGAGATGATGCAGTGTGAGGCGGAGCAGGAGCCG
378 V P F A S R L Y V E M M Q C Q A E Q E P

1261 CTGGTCCGTGTCTTGGTTAATGATCGCGTTGTCCCGCTGCATGGGTGTCCGGTTGATGCT
398 L V R V L V N D R V V P L H G C P V D A

1321 TTGGGGAGATGTACCCGGGATAGCTTTGTGAGGGGGTTGAGCTTTGCTAGATCTGGGGGT
418 L G R C T R D S F V R G L S F A R S G G

1381 GATTGGGCGGAGTGTTTTGCTTAG
438 D W A E C F A *

37/51

tctagaacaataacaggtactccctaggtacccgaaggaccttgtggaaaatgtatggag 60
gtggacacggcaccaccaccacccgcgatggcgacggtggcctaacccttgcctcc 120
ctcaggatggaatccatgtcgactctttaccctcaccatcgccctggatgaaacctccccg 180
ctaagctcacgacgatcgctatttccgaccgatttgaccgcatgggtggagggtgatcc 240
ggtcgatgtccttgccttcatttccggagttcggagacatgaaaggcttatatgaggacgt 300
cccagggtcggggacgaaatccgccctgggctgtgctccttcgctcggaacatctgctgtc 360
cgtgatggctaccatgggctttcttgccattgtgctctccgctcgccctgctctttagaag 420
M G F L A I V L S V A L L F R S 16

gtatgcacccctctacgtccaattctctggygactgacaacgggcgacacatcgggcac 480
T S G T 20

cccggttgggccccggggcaaacatagcgactgcaactcagtcgatcacggctatcaatg 540
P L G P R G K H S D C N S V D H G Y Q C 40

ctttcctgaactctctcataaatggggactctacgcgccctacttctccctccaggacga 600
F P E L S H K W G L Y A P Y F S L Q D E 60

gtctccggttctctgacgtcccagaggactgtcacatcaccttcgtgcaggtgctggc 660
S P F P L D V P E D C H I T F V Q V L A 80

ccgccacggcgcgaggagcccaacccatagcaagaccaaggcgtacgcggcgaccattgc 720
R H G A R S P T H S K T K A Y A A T I A 100

ggccatccagaagagtgccactgcgtttccgggcaaatacgcgttctgcagtcataata 780
A I Q K S A T A F P G K Y A F L Q S Y N 120

ctactccttggactctgaggagctgactcccttcggggcggaaccagctgcgagatctggg 840
Y S L D S E E L T P F G R N Q L R D L G 140

cgcccagttctacgagcgctacaacgccctcaccgcacacatcaacccttcgtccgcgc 900
A Q F Y E R Y N A L T R H I N P F V R A 160

caccgatgcacccgcgtccacgaatccgcccagagaagttcgtcgagggttccaaaccgc 960
T D A S R V H E S A E K F V E G F Q T A 180

tcgacaggacgatcatcacgccaatccccaccagccttcgcctcgcgtggacgtggccat 1020
R Q D D H H A N P H Q P S P R V D V A I 200

ccccgaaggcagcgccctacaacaacacgctggagcacagcctctgcaccgccttcgaatc 1080
P E G S A Y N N T L E H S L C T A F E S 220

cagcaccgctcggcgacgacgcggtcgccaacttcaccgcctgttcgcgcggcgatcgc 1140
S T V G D D A V A N F T A V F A P A I A 240

ccagcgccctggaggccgatcttccggcggtgcagctgtccaccgacgacgtggtaacct 1200
Q R L E A D L P G V Q L S T D D V V N L 260

gatggccatgtgtccgttcgagacggtcagcctgaccgacgacgcgcacacgctgtcgcc 1260
M A M C P F E T V S L T D D A H T L S P 280

gttctgcgacctcttcacggccactgagtgagcagtcacaactacctgctctcgctgga 1320
F C D L F T A T E W T Q Y N Y L L S L D 300

caagtactacggctacggcgggggcaatccgctgggtccggtgcaggggggtcggtgggc 1380
K Y Y G Y G G G N P L G P V Q G V G W A 320

gaacgagctgatggcgcggttaacgcgcgcccccggtgcacgaccacacctgcgtcaacaa 1440
N E L M A R L T R A P V H D H T C V N N 340

caccctcgacgcgagtcgggccaccttcccgctgaacgccacctctacgccgacttctc 1500
T L D A S P A T F P L N A T L Y A D F S 360

38/51

ccacgacagcaacctggtgtcgatcttctgggcgctgggcctgtacaacgycaccgcgcc 1560
 H D S N L V S I F W A L G L Y N G T A P 380
 gctgtcgagacactccgtcgagagcggtctcccagacgggacgggtacgcccgcgcctggac 1620
 L S Q T S V E S V S Q T D G Y A A A W T 400
 ggtgccgttcgcccgtcgcgcgtagctcgagatgatgcagtgtcgcgccgagaaggagcc 1680
 V P F A A R A Y V E M M Q C R A E K E P 420
 gctggtgcgcgtgctggtcaacgaccgggtcatgccgctgcatggctgccctacggacaa 1740
 L V R V L V N D R V M P L H G C P T D K 440
 gctggggcggtgcaagcgggacgctttcgtcgcggggctgagctttgcgagggcgggcg 1800
 L G R C K R D A F V A G L S F A Q A G G 460
 gaactgggcggattgtttctgatgttgagaagaaaggtagatagataggtagtagtacatatg 1860
 N W A D C F 466
 gattgctcggtctctgggtcgttgcacacaatgcatattacgcccgtcaactgccttgcc 1920
 catccacctctcaccctggacgcaaccgagcggtctaccctgcacacggcttccaccgcg 1980
 acgcgacagataaggcgcttttcttacgggggttggggctgggggcagccggagccggag 2040
 agagagaccagcggtgaaaaacgacagaacatagatatcaattcgacgccaattcatgcag 2100
 agtagtatacagacgaactgaaacaaacacatcacttccctcgctcctctcctgtagaag 2160
 acgctcccaccagcgcttctggcccttattcccgtacgctaggtagaccagtcagccag 2220
 acgcatgcctcacaagaacggggcgggggacacactccgctcgtacagcaccacagacg 2280
 tgtacaggaaaaccggcagcgccacaatcgctcgagagccatctgcag 2327

39/51

1	TTCCACGCTGAAAGCCTGACTGCGATTTCCAAGCTGCATGCAGGCTGCTC	50
51	AACTGCCTGCTTATCTTCATCAGACGCAGATACACAACTGGTCTGTAGA	100
101	TGCACCCATGACGGACGAACGCACCGCTCTCTTGGCCTCCAGGGACCCGG	150
151	AGGTCGAGGGCGATGAGGTCCGCCCTCGACGGCCTCCAGTCCCTGTTC	200
201	CAGTTGAGATCTCGCTGCGAACGTGACCGCAGATATGGTTGTCTTCGAC	250
251	GTTTTCTCGCCTTCGAGGAAGAATTGCTGCTGTGACGATGAGTCTGTGT	300
	M S L L L	5
301	TGCTGGTGCTGTCCGGCGGGTTGGTCGCGTTATAGcatgcccccccccc	350
	L V L S G G L V A L Y	16
351	cggtcatattgcccccgctaacgtcttcataactgaagtGTCTCAAGAA	400
	V S R N	20
401	ATCCGCATGTTGATAGCCACTCTTGCAATACAGTGAAGGAGGGTATCAG	450
	P H V D S H S C N T V E G G Y Q	36
451	TGTCGTCCAGAAATCTCCCACTCCTGGGGCCAGTATTCTCCATTCTTCTC	500
	C R P E I S H S W G Q Y S P F F S	53
501	CCTGGCAGACCAGTCCGAGATCTCGCCAGATGTCACAGAACTGCAAGA	550
	L A D Q S E I S P D V P Q N C K I	70
551	TTACGTTTGTCCAGCTGCTTCTCGTCCAGCGCTAGATACCCTACGTCT	600
	T F V Q L L S R H G A R Y P T S	86
601	TCCAAGACCGAGCTGTATTTCGAGCTGATCAGTCGGATTCAAGACGGC	650
	S K T E L Y S Q L I S R I Q K T A	103
651	GACTGCGTACAAAGGCTACTATGCCCTTCTTGAAAGACTACAGATACCAGC	700
	T A Y K G Y Y A F L K D Y R Y Q L	120
701	TGGGAGCGAACGACCTGACGCCCTTTGGGGAAAACCAGATGATCCAGTTG	750
	G A N D L T P F G E N Q M I Q L	136

Fig. 13A

40/51

751	GGCATCAAGTTTTATAACCAATTACAGAGTCTCGCCAGGAATGCCGTCCC G I K F Y N H Y K S L A R N A V P	800 153
801	ATTCTGTTCTGCTCCGGCTCTGATCGGGTCATTGCTCGGGGAGACTTT F V R C S G S D R V I A S G R L F	850 170
851	TCATCGAAGGTTTCCAGAGCGCCAAAGTGCTGGATCCTCATTGAGACAAG I E G F Q S A K V L D P H S D K	900 186
901	CATGACGCTCTCTCCACGATCAACGTGATCATCGAGGAGGGTCCGTCTTA H D A P P T I N V I I E E G P S Y	950 203
951	CAATAACACGCTCGACACCCGGCAGCTGTCCAGTCTTTGAGGACAGCAGCG N N T L D T G S C P V F E D S S G +	1000 220
1001	GGGACATGACCCACAGGAAAAGTTCGCAAGCAATTCGACCAGCTATC G H D A Q E K F A K Q F A P A I	1050 236
1051	CTGGAAAAGATCAAGGACCATCTTCCCGGCGTGGACCTGGCCGTGTCCGA L E K I K D H L P G V D L A V S D	1100 253
1101	TGTACCGTACTTGATGGACTTGTGTCCGTTTGAGACCTTGGCTCGCAACC V P Y L M D L C P F E T L A R N H +	1150 270
1151	<u>ACACAGACACGCTGTCTCCGTTCTGCGCTCTTCCACGCAAGAGGAGTGG</u> T D T L S P F C A L S T Q E E W	1200 286
1201	<u>CAAGCATATGACTACTACCAAGTCTGCGGAAATACTATGGCAATGGCGG</u> Q A Y D Y Y Q S L G K Y Y G N G G	1250 303
1251	GGGTAACCCGTTGGGGCCAGCCCCAAGGCGTGGGGTTTGTCAACGAGTTGA G N P L G P A Q G V G F V N E L I	1300 320
1301	TTGCTCGCATGACCCATAGCCCTGTCCAGGACTACACCACGGTCAACCAC A R M T H S P V Q D Y T T V N H	1350 336
1351	ACTCTTGACTCGAATCCGGCGACATTCCTTTGAACGCGACCGGTGTACGC T L D S N P A T F P L N A T L Y A	1400 353

41/51

1401 AGATTTT CAGCCACGACAACACAATGACGTCAATTTTCGGGGCCTTGGGCC 1450
D F S H D N T M T S I F A A L G L 370

1451 TGTACAACGGGACCGCGAAGCTGTCCACGACCGACATCAAGTCCATTGAA 1500
Y N G T A K L S T T E I K S I E 386
+

1501 GAGACGGACGGCTACTCGGGCGGTGGACCGTTCCGTTCCGGGGGGCGAGC 1550
E T D G Y S A A W T V P F G G R A 403

1551 CTATATCGAGATGATGCAGTGTGATGATTCCGATGAGCCAGTCGTTCCGG 1600
Y I E M M Q C D D S D E P V V R V 420

1601 TGCTGGTCAACGACCGGGTGGTGCCACTGCATGGCTGCGAGGTGGACTCC 1650
L V N D R V V P L H G C E V D S 436

1651 CTGGGGCGATGCAAACGAGACGACTTTGTTCAGGGGACTGAGTTTTCGGCG 1700
L G R C K R D D F V R G L S F A R 453

1701 ACAGGGTGGGAAC TGGGAGGGGTGTTACGCTGCTTCTGAGTAGGTTTATT 1750
Q G G N W E G C Y A A S E 466

1751 CAGCGAGTTTCGACCTTTCTATCCTTCAAACACTGCACAAAGACACACTG 1800

1801 CATGAAATGGTAACAGGCCTGGAGCGTTTTAGAAGGAAAAAAGTT 1845

[illegible]

42/51

Fig. 14A

43/51

CGCCAGTTCTACCAACGGTATCGTGAGCAGGCCGAGAGATTGTGCCATTTGTGCGTGGCAGGCTCCGCGCGAGTCATTCGCTCGGCA 990
R Q F Y H R Y R E Q A R E I V P F V R A A G S A R V I A S A

GAGTTCTTCAACCGGGATTCCAGGATGCCAAGACCGGGATCCAGGTGGAACAAGGACGAGCCIGTGAICAACGIGATCAIT 1080
E F F N R G F Q D A K D R D P R S N K D Q A E P V I N V I I

TCCGAAGAAACIGGCAGTAACAATACCTGGATGGCTGACGTGCCCGCGCGCGAGGACCGACCCAAACCCAGCCGAGAGTTTC 1170
S E E T G S N N T L D G L T C P A A E E A P D P T Q P A E F

CTGCAAGTTTTCGGCCCGGTGCTTGAAAAGATCACTAAACATGCCGGGTGTGAACCTCACCTTGGAGGATGCCCGTTGTTTCATG 1260
L Q V F G P R V L K K I T K H M P G V N L T L E D V P L F M

GATCTTTGTCCGTTTGACACGGTGGGCTCCGACCCAGTTCTTTTCCACGGCAGCTCTCCGTTTGTCTACTTGTTCACGGCCGACGAT 1350
D L C P F D T V G S D P V L F P R Q L S P F C H L F T A D D

TGGATGGCCTACGATTACTACTACACCTCGACAAATACTACAGCCACGGGGGCGGAGCGCATTTGGCCCGTCCCGCGGGTCCGGTTC 1440
W M A Y D Y Y T L L D K Y Y S H G G G S A F G P S R G V G F

GTCAACGAGCTGATTCGGGTATGACGGGAAATCTTCCGTCAGGACCCACACAGTCACACACTCTCGATGACAAACCCGGAACCT 1530
V N E L I A R M T G N L P V K D H T T V N H T L D D N P E T

TTCCCGTTGGACCGCTGCTCTACGAGACTTTTGGCAGCACAACCATGACGGGCATCTTTTCCGCAATGGCCCTGTACAACCGGCACA 1620
F P L D A V L Y A D F S H D N T M T G I F S A M G L Y N G T

AAGCCCGCTGTCAGCTCCAAGATTACGCCCTCCGACGGGTGCAGCGGATGGATATGCGGCATCGTGGACGGTCCCGTTCCGACGGAGG 1710
K P L S T S K I Q P P T G A A A D G Y A A S W T V P F A A R

Fig. 14B

45/51

AAGCTTGGGCAAACCTCATCATGCTCATCTTGATGATTCCACTGTTTCAGCTACCTGGCTGCTGCTTCTCTGTTGGGTTTCATC
80
HindIII M L I L M I P L F S Y L A A A S L

CTTTGCCCCTGTCTCGATGTTAAATACTAAACATATTTACCAGACGTGTACTCTCCCTCAGCCAGTGTCTGTGACA
160

R V L S P Q P V S C D

GCCCGGAGCTTGGTTACCAATGCGACCAGCAGACAACGCACACCTGGGGTCAATACTCACCCTTCTTCTGTCTCCGTC
240
S P E L G Y Q C D Q Q T T H T W G Q Y S P F F S V P S

GAGATCTCCCTTCCGTTCTGATGGCTGCCGCTCACCTTCGCCCCAAGTTCTCTCCCGCCACGGCGCCGCTTCCCAAC
320
E I S P S V P D G C R L T F A Q V L S R H G A R F P T

CCCGGGTAAAGCCGCGCCATCTCCGCTGTCTCACCAAAATCAAAACCTCTGCCACCTGGTACGGTTCCGACTTTTCAGT
400
P G K A A A I S A V L T K I K T S A T W Y G S D F Q

TCATCAAGAACTACGACTATGTACTTGGCGTAGACCACCTGACCGCTTCGGCGAGCAAGAAATGGTCAACTCCGGCATC
480
F I K N Y D Y V L G V D H L T A F G E Q E M V N S G I

AAGTTCTACCAGCGCTACTCCTCCCTCATCCAGACAGAAGACTCGGATACGCTCCCTTCGTCCGCGCCTCTGGCCAGGA
560
K F Y Q R Y S S L I Q T E D S D T L P F V R A S G Q E

ACGCGTCATCGCCTCCGCGGAGAACTTACCACCGGCTTCTACTCGGCCCTCTCAGCCGACAAGAACCCTCCTTCTCTCT
640
R V I A S A E N F T T G F Y S A L S A D K N P P S S

TACCAAGACCAGAAATGGTCATCATTTCTGAGGAGCCAACAGCCAACAACACCATGCACCACGGCCTCTGCCGCTCCTTT
720
L P R P E M V I I S E E P T A N N T M H H G L C R S F

GAAGATTCCACCACCGGCGACCAAGCCCAAGCGGAATTCATCGCCGCCACCTTCCACCCATCACCGCCCGTCTCAACGC
800
E D S T T G D Q A Q A E F I A A T F P P I T A R L N A

CCAAGGTTTCAAAGGCGTCACCCTCTCCAACACCGACGTCCTATCACTAATGGACCTCTGCCCCCTTGACACCGTCGCCT
880
Q G F K G V T L S N T D V L S L M D L C P F D T V A

ACCCCTTTCTCTCCCTCACCACCACCTCTTCCGTTTCTGGAGGCGGCAAGTTATCCCCCTTCTGCTCTCTTTTCACTGCC
960
Y P L S S L T T T S S V S G G G K L S P F C S L F T A

AGCGACTGGACAATCTACGATTACCTCCAGTCCCTAGGGAAATACTACGGTTTCGGCCCCGGTAATTCCTAGCTGCCAC
1040
S D W T I Y D Y L Q S L G K Y Y G F G P G N S L A A T

CCAGGGGGTAGGGTACGTCAACGAGCTTATCGCCCGCTTGATCCGTGCTCCCGTCGTAGATCACACGACGACCAACTCTA
1120
Q G V G Y V N E L I A R L I R A P V V D H T T T N S

CTCTTGATGGCGACGAAAAACGTTTCCGTTGAACAGAACGGTGTATGCGGATTTTCCCATGATAATGATATGATGAAT
1200
T L D G D E K T F P L N R T V Y A D F S H D N D M M N

46/51

ATCCTGACTGCTTTGCGGATATTCGAGCATATCAGTCCGATGGATAACACCACTATCCCGACCACTATGGCCAGACAGG
1280
I L T A L R I F E H I S P M D N T T I P T N Y G Q T G

AGATGACGGGGTGAAGGAAAGGGATTGTTC AAGGTTAGTTGGGCGGTGCCCTTTGCTGGGAGGGTGTACTTTGAGAAA
1360
D D G V K E R D L F K V S W A V P F A G R V Y F E K

TGGTTTGTGATGCCGATGGGGATGGCAAGATTGATAGTGATGAGGCTCAGAAAGAGTTGGTGAGGATTTGGTTAATGAT
1400
M V C D A D G D G K I D S D E A Q K E L V R I L V N D

CGGGTGATGAGATTGAATGGGTGTGATGCTGATGAACAGGGTAGGTGTGGATTGGAGAAGTTTGTGGAGAGTATGGAGTT
1520
R V M R L N G C D A D E Q G R C G L E K F V E S M E F

TGCGAGGAGAGGGGGGAGTGGGAGGAGAGGTGTTTTGTTTAGCTCTAGA
A R R G G E W E E R C F V XbaI

47/51

```

1 ..MLILMIPLFYSYLAASLRVLSPQPVSCDSPELGYQCDQQTHTTWGQYS 48
| :| :| :| :| | | | . | . | | :| :| :| :| ...| | | | |
1 MTGLGVMMVVMVGFLAIASLQSESR...PCDTPDLGFQCQGTASHFWGQYS 47
.
49 PFFSVPSSEISPSVPDGCRLTFAQVLSRHGARFPTPGKAAAISAVLTKIKT 98
| :| | | | | :| :| :| :| | :| | | | | | | | | | :| | . . . :| .
48 PYFSVPSELDAIPDDCEVTFAQVLSRHGARAPTLKRAASYVDLIDRIHH 97
.
99 SATWYGSDFOFIKNYDYVLGVDHLTAFGGEQEMVNNGIKFYORYSSLIQTE 148
:| . | | :| :| :| :| . | | . | | . | | | :| :| :| | | | | . | | . |
98 GAISYGPGEFLRTYDYTLGADELTRTGQQQMVNNGIKFYRRYRAL.... 143
.
149 DSDTLFPVRASGOERVIASAENFTTGFYSALSADKNPPSSLPRP.EMVII 197
.....:| | | | | :| :| :| | | | | | | | | | | :| :| . . . | :| :| :|
144 ARKSIFPVRTAGQDRVVHSAENFTQGFGHSALLADRGSTVRPTLPYDMVVI 193
.
198 SEEPTANNTMHGCLCRSFED...STTGDQAQAEFIAATFPPITARLNAQG 244
.| :| . | | | :| :| :| | . | | :| | . | | :| | . . . . . | | | | :| | .
194 PETAGANNTLHNDLCTAFEEGPYSTIGDDAQDTYLSTFAGPITARVNA.N 242
.
245 FKGVTLSNTDVLSLMDLCPFDTVAYPLSSLTTTSSVSXGGGK.LSPFCSLF 293
:| . | . | . . . | . . | | | | :| | | | . | . | . . . :| :| :| | | | | |
243 LPGANLTDADTVALMDLCPFETVASSSSDPATADAGGGNGRPLSPFCRLF 292
.
294 TASDWTIYDYQLSLGKYGYGFGPNGLAATQGVGYVNELIARLIRAPVVDH 343
..| :| | | | | :| :| :| :| | | | . :| :| | | | :| | | :| | | | | . | | |
293 SESEWRAYDYQLSVGKWYGYGPGNPLGPTQGVGVFNELLARLAGVPVRDG 342
.
344 TTTNSTLDGDEKTFPLNRTVYADFSDNDMMNILTALRIFEHISPMDNTT 393
|. | | . | | | | :| | | | . :| | | | | | | | | . :| . | | :| :| :| :|
343 TSTNRTLGDGDPRTFPLGRPLYADFSDNDMMGVLGALGAYDGVPPLD... 389
.
394 IPTNYGQTGDDGVKERDLFKVSWAVPFAGRIFYEKMVCADAGDGKIDS.D 442
.:| :| . . | :| :| . | | | | :| :| :| | | | | | :| :| :| :| :|
390 .....KTARRDPEELGGYAASWAVPFARAIIYVEKMRCSGGGGGGGGGEG 433
.
443 ..EAQKELVRILVNDRVMLNGCDADEQGRCGLEKFVESMEFARRGGWE 490
| :| :| :| :| | | | | | . | | :| | | . | . | :| :| :| | | | . | . | :|
434 RQEKDEEMVRVLVNDRVMTLKGCCADERGMCTLERFIESMAFARGNGKWD 483
.
491 ERCFV 495
||.
484 L.CFA 487

```

48/51

Peniophora numbers	1				37
Alignment numbers	1				50
P_involtus_A1ML	FGFVALACLL	SLSEVLATSV	P.....KNT	APTFPIPESE
P_involtus_A2MH	LGFVTLACLI	HLSEVFAASV	P.....RNI	APKFSIPESE
T_pubescens	MAFSILASLL	FVCYAYARAV	PRAHIPLRDT	SACLDVTRDV
A_pediades	MSLFIGGCLL	VFLQASAYGG	VVQATFVQPFFPPQI
P_lyciiMV	SSAFAPSILL	SLMSSLALST	QFSF.....V	AAQLPIPAQN
A_fumigatusMVTL	TFLLSAAYLL	.SGRVSAAPS	SAGSKSCDTV	DLGYQCSPAT
consphyAMGVF	VVLLSIATLF	GSTSGTALGP	RGNSHSCDTV	DGGYQCFPEI
A_nidulansMAFF	TVALSLYLL	..SRVSAQAP	VVQNHSCNTA	DGGYQCFPNV
A_ficuum_NRRL3135MGVS	AVLLPLYLLS	GVTSGLAIPA	SRNQSSCDTV	DQGYQCFSET
A_terreusMGFL	AIVLSVALLF	RSTSGTPLGP	RGKHSDCNSV	DHGYQCFPEL
T_thermoMSLL	LLVLSGGLVA	LYVS...RNP	HVDSHSCNTV	EGGYQCRPEI
T_lanuginosa		MAGIGLGSFL	VLLLQFSALL	TASPAIPFW	RKKHPNVD..
M_thermophilaMTGL	GVMVMVGFL	AIASL.....	QSESRPCDTP	DLGFQCGTAI
C_foecundissimumML	ILMIPLFSYL	AAASL	RVLSPSCDSP	ELGYQCDQQT
				QPV	
	38				83
	51				100
P_involtus_A1	QRNWSPPSPY	FPLAEYKA..	..PPAGCQIN	QVNIQHRGA	RFPTSGATTR
P_involtus_A2	QRNWSPPSPY	FPLAEYKA..	..PPAGCEIN	QVNIQHRGA	RFPTSGAATR
T_pubescens	QQSWSMSPY	FPAATYVA..	..PPASCQIN	QVHIIQHRGA	RFPTSGAAKR
A_pediades	QDSWAAATPY	YPVQAYTP..	..PPKCKIT	QVNIQHRGA	RFPTSGAGTR
P_lycii	TSNWGPYDPF	FPVEPYAA..	..PPEGCTVT	QVNIQHRGA	RWPTSGARSR
A_fumigatus	SHLWGQYSPF	FSLEDELSVS	SKLPKDCRIT	LVQVLSRGA	RYPTSSKSKK
consphyA	SHLWGQYSPY	FSLEDESAIS	PDVPPDCRVT	FVQVLSRGA	RYPTSSKSKA
A_nidulans	SHVWGQYSPY	FSIEQESAIS	EDVPHGCEVT	FVQVLSRGA	RYPTSSKSKA
A_ficuum_NRRL3135	SHLWGQYAPF	FSLANESVIS	PEVPAGCRVT	FAQVLSRGA	RYPTDSKGGK
A_terreus	SHKWGLYAPY	FSLQDESPFP	LDVPEDCHIT	FVQVLSRGA	RSPTHSKTKA
T_thermo	SHSWGQYSPF	FSLADQSEIS	PDVPQCKIT	FVQVLSRGA	RYPTSSKTEL
T_lanuginosa	ARHWGQYSPF	FSLAEVSEIS	PAVPKGCVE	FVQVLSRGA	RYPTAHKSEV
M_thermophila	SHFWGQYSPY	FSVP..SELD	ASIPDDCEVT	FAQVLSRGA	RAPTLKRAAS
	THTWGQYSPF	FSVP SEIS	PSVPDGCRLT	FAQVLSRGA	RFPTPGKAAA
	84				133
	101				150
P_involtus_A1	IKAGLTKLQG	VQNFDAKFN	FIKSFYDLG	NSDLVPFGAA	QSFDAGQEA
P_involtus_A2	IKAGLSKLQS	VQNFDPKFD	FIKSFTYDLG	TSDLVPFGAA	QSFDAGLEV
T_pubescens	IQTAVAKLKA	ASNYTDPILA	FVTNYTSLG	QDSLVELGAT	QSSEAGQEA
A_pediades	IQAAVKKLQS	AKTYTDPRLD	FLTNYTYTLG	HDDLVPFGAL	QSSQAGEET
P_lycii	QVA AVAKIQM	ARPFDPKYE	FLNDFVYKFG	VADLLPFGAN	QSHQTGTDMY
A_fumigatus	YKKLVTAIQA	NATDFKKGFA	FLKTYNYTLG	ADDLTPFGEQ	QLVNSGIKFY
consphyA	YSALIEAIQK	NATAFKGKYA	FLKTYNYTLG	ADDLTPFGEN	QMVNSGIKFY
A_nidulans	YSGLIEAIQK	NATSWFGQYA	FLESYNYTLG	ADDLTFGEN	QMVDSGAKFY
A_ficuum_NRRL3135	YSALIEEIQQ	NATTFDGKYA	FLKTYNYSLG	ADDLTPFGEQ	ELVNSGIKFY
A_terreus	YAATIAAIQK	SATAFPKGYA	FLQSYNYSLD	SEELTPFGRN	QLRDLGAQFY
T_thermo	YSQLISRIQK	TATAYKGYA	FLKDYRYQLG	ANDLTPFGEN	QMIQLGIKFY
T_lanuginosa	YAELLQRIQD	TATEFKGDF	FLRDYAYHLG	ADNLTRFGEE	QMMESGRQFY
M_thermophila	YVDLIDRIH	GAISYGPGE	FLRTYDYTLG	ADELTRTGQQ	QMVNSGIKFY
	ISAVLTKIKT	SATWYGSDFO	FIKNYDYVLG	VDHLTAFGEQ	EMVNSGIKFY
	134				176
	151				200
P_involtus_A1	ARYSKLVSKN	NLPFIRADGS	DRVVDSATNW	TAGFASA...SHNTVQ
P_involtus_A2	ARYSKLVSSD	NLPFIRSDGS	DRVVDATNW	TAGFASA...SRNAIQ

49/51

T_pubescens	TRYSSLVSAD	ELPFVRASGS	DRVATANNW	TAGFALA...	SSNSIT
A_pediades	QRYSLVSKE	NLPFVRASS	NRVDSATNW	TEGFSAA...	SHHVLN
P_lycii	TRYSTLFEGG	DVPFVRAAGD	QRVDSSTNW	TAGFGDA...	SGETVL
A_fumigatus	QRYKAL.ARS	VVPFIRASGS	DRVIASGEKF	IEGFQQAkla	DPGA.TNRAA
consphyA	RRYKAL.ARK	IVPFIRASGS	DRVIASAEKF	IEGFQSAKLA	DPGSQPHQAS
A_nidulans	RRYKNL.ARK	NTPFIRASGS	DRVVASAEKF	INGFRKAQLH	DHGS..KRAT
A_ficuum_NRRL3135	QRYESL.TRN	IVPFIRSSGS	SRVIASGKKF	IEGFQSTKLG	DPRAQPGQSS
A_terreus	ERYNAL.TRH	INPFVRATDA	SRVHESAIEKF	VEGFQTARQD	DHHANPHQPS
T_thermo	NHYKSL.ARN	AVPFVRCSSGS	DRVIASGRLE	IEGFQSAKVL	DPHSDKHDAP
T_lanuginosa	HRYREQ.ARE	IVPFVRAAGS	ARVIASAEFF	NRGFQDAKDR	DPRSNKDQAE
M_thermophila	RRYRAL.ARK	SIPFVRTAGQ	DRVVHSAENF	TQGFHSALLA	DRGSTVRPTL
	QRYSSLIDSD	TLPFVRASGQ	ERVIASAENF	TTGFYSALSA	DKNPPSSLPR
	QTE				

177

217

201

250

P_involtus_A1	PKLNLILPQT	G..NDTLEDN	MCPAAGD...	SDPQVNA	WLAVAFPSIT
P_involtus_A2	PKLDLILPQT	G..NDTLEDN	MCPAAGE...	SDPQVDA	WLASAFPSVT
T_pubescens	PVLSVISEA	G..NDTLDDN	MCPAAGD...	SDPQVNO	WLAQFAPPMT
A_pediades	PILFVILSES	L..NDTLDDA	MCPNAGS...	SDPQTGI	WTSIYGTPIA
P_lycii	PTLQVVLQEE	G..NCTLCNN	MCPNEVD...	GD..ESTT	WLGVFAPNIT
A_fumigatus	PAISVPIES	ETFNNTLDHG	VCTKFEA...	SQLGDEVAAN	FTALFAPDIR
consphyA	PVIDVPIEG	SGYNNNTLDHG	TCTAFED...	SELGDDVEAN	FTALFAPAIR
A_nidulans	PVVNVPIEI	DGFNNNTLDHS	TCVSFEN...	DERADEIEAN	FTAIMGPPIR
A_ficuum_NRRL3135	PKIDVWISEA	SSSNNTLDPG	TCTVFED...	SELADTVEAN	FTATFVPSIR
A_terreus	PRVDVAIPEG	SAYNNTLEHS	LCTAFES...	STVGDDAVAN	FTAVFAPAIA
T_thermo	PTINVIIIEG	PSYNNNTLDTG	SCPVFED...	SSGGHDAQEK	FAKQFAPAIL
T_lanuginosa	PVINVIIEE	TGSNNNTLDGL	TCPAEEE...	AP..DPTQPAE	FLQVFGPRVL
M_thermophila	PYDMVVIPE	AGANNNTLHND	LCTAFEEGPY	STIGDDAQDT	YLSTFAGPIT
	P.EMVIIEE	PTANNNTMHG	LCRSFED	STTGDDAQAE	FIAATFFPIT

218

252

251

300

P_involtus_A1	ARLNAAAPSV	NLTDTDAFNL	VSLCAFLTVS	KEKK.....	S
P_involtus_A2	AQLNAAAPGA	NLTDAADFNL	VSLCPFMFVS	KEQK.....	S
T_pubescens	ARLNAGAPGA	NLTDTDTYNL	LTLCPFETVA	TERR.....	S
A_pediades	NRLNQAPGA	NITAADVSNL	IPLCAFETIV	KETP.....	S
P_lycii	ARLNAAAPSA	NLSDSDALTL	MDMCPFDTL	SGNA.....	S
A_fumigatus	ARAEXHLPV	TLTDEDVVS	MDMCSFDTVA	RTSD..ASQ.	LS
consphyA	ARLEADLPV	TLTDEDVVYL	MDMCPFETVA	RTSD..ATE.	LS
A_nidulans	KRLENDLPGI	KLTNENVIYL	MDMCSFDTMA	RTAH..GTE.	LS
A_ficuum_NRRL3135	QRLENDLSGV	TLTDTTEVTYL	MDMCSFDTIS	TSTV..DTK.	LS
A_terreus	QRLEADLPV	QLSTDDVVNL	MAMCPFETVS	LTDD..AHT.	LS
T_thermo	EKIKDHLPGV	DLAVSDVPYL	MDLCPFETLA	RNHT..DT..	LS
T_lanuginosa	KKITKHMPGV	NLTLEDVPLF	MDLCPFDTVG	SDPVLFPQR.	LS
M_thermophila	ARVNANLPGA	NLTDAADTVAL	MDLCPFETVA	SSSDPATAD	AGGGNGRPLS
	ARLNAGFKGV	TLSTNTDVLSL	MDLCPFDTVG	YPLSSLTTS	SVSGGGK LS

Q

253

300

301

350

P_involtus_A1	DFCTLFEGIP	GSFEAFAYGG	DLDKIFYGTGY	GOELGPVQGV	GYVNELIARL
P_involtus_A2	DFCTLFEGIP	GSFEAFAYAG	DLDKIFYGTGY	GOALGPVQGV	GYINELLARL
T_pubescens	EFCDIYEELQ	AE.DAFAYNA	DLDKIFYGTGY	GQPLGPVQGV	GYINELIARL
A_pediades	PFCNLFT..P	EEFAQFEYFG	DLDKIFYGTGY	GQPLGPVQGV	GYINELLARL
P_lycii	PFCDLFT..A	EEYVSYYEYY	DLDKYYGTGP	GNALGPVQGV	GYVNELIARL
A_fumigatus	PFCQLFT..H	NEWKKYNYLQ	SLGKYYGYGA	GNPLGPAQGI	GFTNELIARL

Fig 17B

50/51

consphyA	PFCALFT..H	DEWROYDYLO	SLGKYYGYGA	GNPLGPAQGV	GFANELIARL
A_nidulans	PFCAIFT..E	KEWLOYDYLO	SLSKYYGYGA	GSPLGPAQGI	GFTNELIARL
A_ficuum_NRR3135	PFCDLFT..H	DEWINYDYLO	SLKKYYGHGA	GNPLGPTQGV	GYANELIARL
A_terreus	PFCDLFT..A	TEWTQYNYLL	SLDKYYGYGG	GNPLGPVQGV	GWANELMARL
T_thermo	PFCALST..Q	EEWQAYDYYQ	SLGKYYGNNG	GNPLGPAQGV	GFVNELIARM
T_lanuginosa	PFCHLFT..A	DDWMAYDYYY	TLDKYYSHGG	GSAFGPSRGV	GFVNELIARM
M_thermophila	PFCRLFS..E	SEWRAYDYLO	SVGKYYGYGP	GNPLGPTQGV	GFVNELLARL
	PFCSLFT A	SDWTIYDYLO	SLGKYYGFGP	GNSLAATQGV	GYVNELIARL

301

349

351

400

P_involtus_A1	TNS.AVRDNT	QTNRTLDASP	VTFPLNKTIFY	ADFSHDNLMV	AVFSAMGLFR
P_involtus_A2	TNS.AVNDNT	QTNRTLDAAP	DTFPLNKTMY	ADFSHDNLMV	AVFSAMGLFR
T_pubescens	TAQ.NVSDHT	QTNSTLDSSP	ETFPLNRTLY	ADFSHDNQMV	AIFSAMGLFN
A_pediades	TEM.PVRDNT	QTNRTLDSSP	LTFPLDRSIY	ADLSHDNQMI	AIFSAMGLFN
P_lycii	TGQ.AVRDET	QTNRTLDSDP	ATFPLNRTFY	ADFSHDNTMV	PIFAALGLFN
A_fumigatus	TRS.PVQDHT	STNSTLVSNP	ATFPLNATMY	VDFSHDNSMV	SIFFALGLYN
consphyA	TRS.PVQDHT	STNHTLDSNP	ATFPLNATLY	ADFSHDNSMI	SIFFALGLYN
A_nidulans	TQS.PVQDNT	STNHTLDSNP	ATFPLDRKLY	ADFSHDNSMI	SIFFAMGLYN
A_ficuum_NRR3135	THS.PVHDDT	SSNHTLDSSP	ATFPLKSTLY	ADFSHDNGII	SILFALGLYN
A_terreus	TRA.PVHDHT	CVNNTLDASP	ATFPLNATLY	ADFSHDSNLV	SIFWALGLYN
T_thermo	THS.PVQDYT	TVNHTLDSNP	ATFPLNATLY	ADFSHDNTMT	SIFAALGLYN
T_lanuginosa	TGNLPVKDHT	TVNHTLDDNP	ETFPLDAVLY	ADFSHDNTMT	GIFSAMGLYN
M_thermophila	A.GVPVRDGT	STNRTLDGDP	RTFPLGRPLY	ADFSHDNDMM	GVLGALGAYD
	I RAPVVDHT	TTNSTLDGDE	KTFPLNRTVY	ADFSHDNDMM	NILTALRIFE

350

383

401

450

P_involtus_A1	QPAPLSTSV	NPWR.....T	WRTSSLVPFS	GRMVVERLSC
P_involtus_A2	QSAPLSTSTP	DPNR.....T	WLTSSVVPFS	ARMAVERLSC
T_pubescens	QSAPLDPTTP	DPAR.....T	FLVKKIVPFS	ARMVVERLDC
A_pediades	QSSPLDPSFP	NPKR.....T	WVTSRLTPFS	ARMVTERLLC	QRDGTGSGGP
P_lycii	ATA.LDPLKP	DENR.....L	WVDSKLVVFS	GHMTVEKLAC
A_fumigatus	GTEPLSRTSV	ESAKE..LDG	YSASWVVPFG	ARAYFETMQC
consphyA	GTAPLSTTSV	ESIEE..TDG	YSASWTVVFP	ARAYVEMMQC
A_nidulans	GTQPLSMDSV	ESIQE..MDG	YAASWTVVFP	ARAYFELMQC
A_ficuum_NRR3135	GTKPLSTTTV	ENITQ..TDG	FSSAWTVVFP	SRLYVEMMQC
A_terreus	GTAPLSQTSV	ESVSQ..TDG	YAAAWTVVFP	ARAYVEMMQC
T_thermo	GTAKLSTTEI	KSIEE..TDG	YSAAWTVVFP	GRAYIEMMQC
T_lanuginosa	GTKPLSTSKI	QPPTGAAADG	YAASWTVVFP	ARAYVELLRC	ETETSSEEEE
M_thermophila	GVPPLDKTAR	RDPEE..LGG	YAASWAVPFA	ARIYVEKMRC	SGGGGGGGGG
	HISPMDQTGD	DGVKE RDL	FKVSWAVPFA	GRVYFEKMVC	DADGDGKIDS

NTTIPTNYG

384

425

451

500

P_involtus_A1FGT	TKVRVLVDQ	VQPLEFCGGD	RNGLCTLAKF	VESQTFARSD
P_involtus_A2AGT	TKVRVLVDQ	VQPLEFCGGD	QDGLCALDKF	VESQAYARSG
T_pubescensGGA	QSVRLLVNDA	VQPLAFCGAD	TSGVCTLDFA	VESQAYARND
A_pediades	SRIMRNGNVQ	TFVRILVNDA	LQPLKFCGGD	MDSLCTLEAF	VESQKYARED
P_lyciiSGK	EAVRVLVNDA	VQPLEFCGG.	VDGVCELSAF	VESQTYAREN
A_fumigatus	K..S...EKE	PLVRALINDR	VVPLHGCDVD	KLGRCKLNDF	VKGLSWARSG
consphyA	Q..A...EKE	PLVRVLVNDR	VVPLHGCAVD	KLGRCKRDDF	VEGLSFARSG
A_nidulans	E.....KKE	PLVRVLVNDR	VVPLHGCAVD	KFGRCTLDDW	VEGLNFARSG
A_ficuum_NRR3135	Q..A...EQA	PLVRVLVNDR	VVPLHGCPVD	ALGRCTRDSF	VRGLSFARSG
A_terreus	R..A...EKE	PLVRVLVNDR	VMPLHGCPVD	KLGRCKRDAF	VAGLSFAQAG
T_thermo	D..D...SDE	PVVRVLVNDR	VVPLHGCEVD	SLGRCKRDDF	VRGLSFARQG

51/51

T_lanuginosa E..G...EDE PFVRVLVND R VVPLHGCRVD RWGRCRRDEW IKELTFARQG
M_thermophila E..GRQEKDE EMVRVLVND R VMTLKGCAD ERGMCTLERF IESMAFARGN
D EAQK ELVRILVND R VMRLNGCDAD EQGRCGLEKF VESMEFARRG

426 439

501 514

P_involtus_A1 GAGDFEKCFA TSA.
P_involtus_A2 GAGDFEKCFA TTV.
T_pubescens GEGDFEKCFA T...
A_pediades GQGDFFKCFD
P_lycii GQGDFAKCGF VPSE
A_fumigatus ..GNWGECSF
consphyA ..GNWAECSF *...
A_nidulans ..GNWKTCTF L...
A_ficum_NRR13135 ..GDWAECSF
A_terreus ..GNWADCF.
T_thermo ..GNWEGCYA ASE.
T_lanuginosa ..GHWDRCF.
M_thermophila ..GKWDLCFA
GEWEECFV
R

SEQUENCE LISTING

<212> DNA

<213> Cladorrhinum foecundissimum

5 <220>

<221> intron

<222> (71)..(126)

<220>

10 <221> CDS

<222> (20)..(70)

<220>

<221> CDS

15 <222> (127)..(1563)

<220>

<221> sig_peptide

<222> (20)..(64)

20

<400> 1

aagcttgggc aaactcatc atg ctc atc ttg atg att cca ctg ttc agc tac 52
 Met Leu Ile Leu Met Ile Pro Leu Phe Ser Tyr
 1 5 10

25

ctg gct gct gct tct ctg tgggttcac ctttgcccct gtctcgatgt 100
 Leu Ala Ala Ala Ser Leu
 15

30 taaaatacta aacatatttc accaga cgt gta ctc tcc cct cag cca gtg tcc 153
 Arg Val Leu Ser Pro Gln Pro Val Ser
 20 25

tgt gac agc ccg gag ctt ggt tac caa tgc gac cag cag aca acg cac 201
 35 Cys Asp Ser Pro Glu Leu Gly Tyr Gln Cys Asp Gln Gln Thr Thr His
 30 35 40

acc tgg ggt caa tac tca ccc ttc ttc tct gtc ccg tca gag atc tcc 249
 Thr Trp Gly Gln Tyr Ser Pro Phe Phe Ser Val Pro Ser Glu Ile Ser
 40 45 50 55

cct tcc gtt cct gat ggc tgc cgc ctc acc ttc gcc caa gtt ctc tcc 297
 Pro Ser Val Pro Asp Gly Cys Arg Leu Thr Phe Ala Gln Val Leu Ser
 60 65 70

45

cgc cac ggc gcc cgc ttc cca acc ccg ggt aaa gcc gcc gcc atc tcc 345
 Arg His Gly Ala Arg Phe Pro Thr Pro Gly Lys Ala Ala Ala Ile Ser
 75 80 85 90

50 gct gtc ctc acc aaa atc aaa acc tct gcc acc tgg tac ggt tcc gac 393
 Ala Val Leu Thr Lys Ile Lys Thr Ser Ala Thr Trp Tyr Gly Ser Asp
 95 100 105

ttt cag ttc atc aag aac tac gac tat gta ctt ggc gta gac cac ctg 441
 55 Phe Gln Phe Ile Lys Asn Tyr Asp Tyr Val Leu Gly Val Asp His Leu
 110 115 120

	acc gcg ttc ggc gag caa gaa atg gtc aac tcc ggc atc aag ttc tac	489
	Thr Ala Phe Gly Glu Gln Glu Met Val Asn Ser Gly Ile Lys Phe Tyr	
	125 130 135	
5	cag cgc tac tcc tcc ctc atc cag aca gaa gac tcg gat acg ctc ccc	537
	Gln Arg Tyr Ser Ser Leu Ile Gln Thr Glu Asp Ser Asp Thr Leu Pro	
	140 145 150	
10	ttc gtc cgc gcc tct ggc cag gaa cgc gtc atc gcc tcc gcc gag aac	585
	Phe Val Arg Ala Ser Gly Gln Glu Arg Val Ile Ala Ser Ala Glu Asn	
	155 160 165 170	
15	ttc acc acc ggc ttc tac tcg gcc ctc tca gcc gac aag aac cct cct	633
	Phe Thr Thr Gly Phe Tyr Ser Ala Leu Ser Ala Asp Lys Asn Pro Pro	
	175 180 185	
20	tcc tcc tta cca aga cca gaa atg gtc atc att tct gag gag cca aca	681
	Ser Ser Leu Pro Arg Pro Glu Met Val Ile Ile Ser Glu Glu Pro Thr	
	190 195 200	
25	gcc aac aac acc atg cac cac ggc ctc tgc cgc tcc ttt gaa gat tcc	729
	Ala Asn Asn Thr Met His His Gly Leu Cys Arg Ser Phe Glu Asp Ser	
	205 210 215	
30	acc acc ggc gac caa gcc caa gcg gaa ttc atc gcc gcc acc ttc cca	777
	Thr Thr Gly Asp Gln Ala Gln Ala Glu Phe Ile Ala Ala Thr Phe Pro	
	220 225 230	
35	ccc atc acc gcc cgt ctc aac gcc caa ggt ttc aaa ggc gtc acc ctc	825
	Pro Ile Thr Ala Arg Leu Asn Ala Gln Gly Phe Lys Gly Val Thr Leu	
	235 240 245 250	
40	tcc aac acc gac gtc cta tca cta atg gac ctc tgc ccc ttt gac acc	873
	Ser Asn Thr Asp Val Leu Ser Leu Met Asp Leu Cys Pro Phe Asp Thr	
	255 260 265	
45	gtc gcc tac ccc ctt tcc tcc ctc acc acc acc tct tcc gtt tct gga	921
	Val Ala Tyr Pro Leu Ser Ser Leu Thr Thr Thr Ser Ser Val Ser Gly	
	270 275 280	
50	ggc ggc aag tta tcc ccc ttc tgc tct ctt ttc act gcc agc gac tgg	969
	Gly Gly Lys Leu Ser Pro Phe Cys Ser Leu Phe Thr Ala Ser Asp Trp	
	285 290 295	
55	aca atc tac gat tac ctc cag tcc cta ggg aaa tac tac ggt ttc ggc	1017
	Thr Ile Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly Phe Gly	
	300 305 310	
60	ccc ggt aat tcc cta gct gcc acc cag ggg gta ggg tac gtc aac gag	1065
	Pro Gly Asn Ser Leu Ala Ala Thr Gln Gly Val Gly Tyr Val Asn Glu	
	315 320 325 330	
65	ctt atc gcc cgc ttg atc cgt gct ccc gtc gta gat cac acg acg acc	1113
	Leu Ile Ala Arg Leu Ile Arg Ala Pro Val Val Asp His Thr Thr Thr	
	335 340 345	

aac tct act ctt gat ggc gac gaa aaa acg ttt ccg ttg aac aga acg 1161
 Asn Ser Thr Leu Asp Gly Asp Glu Lys Thr Phe Pro Leu Asn Arg Thr
 350 355 360

5

gtg tat gcg gat ttt tcc cat gat aat gat atg atg aat atc ctg act 1209
 Val Tyr Ala Asp Phe Ser His Asp Asn Asp Met Met Asn Ile Leu Thr
 365 370 375

10

gct ttg cgg ata ttc gag cat atc agt ccg atg gat aac acc act atc 1257
 Ala Leu Arg Ile Phe Glu His Ile Ser Pro Met Asp Asn Thr Thr Ile
 380 385 390

15

ccg acc aac tat ggc cag aca gga gat gac ggg gtg aag gaa agg gat 1305
 Pro Thr Asn Tyr Gly Gln Thr Gly Asp Asp Gly Val Lys Glu Arg Asp
 395 400 405 410

20

ttg ttc aag gtt agt tgg gcg gtg ccc ttt gct ggg agg gtg tac ttt 1353
 Leu Phe Lys Val Ser Trp Ala Val Pro Phe Ala Gly Arg Val Tyr Phe
 415 420 425

25

gag aaa atg gtt tgt gat gcg gat ggg gat ggc aag att gat agt gat 1401
 Glu Lys Met Val Cys Asp Ala Asp Gly Asp Gly Lys Ile Asp Ser Asp
 430 435 440

30

gag gct cag aaa gag ttg gtg agg att ttg gtt aat gat ccg gtg atg 1449
 Glu Ala Gln Lys Glu Leu Val Arg Ile Leu Val Asn Asp Arg Val Met
 445 450 455

35

aga ttg aat ggg tgt gat gct gat gaa cag ggt agg tgt gga ttg gag 1497
 Arg Leu Asn Gly Cys Asp Ala Asp Glu Gln Gly Arg Cys Gly Leu Glu
 460 465 470

40

aag ttt gtg gag agt atg gag ttt gcg agg aga ggg ggg gag tgg gag 1545
 Lys Phe Val Glu Ser Met Glu Phe Ala Arg Arg Gly Gly Glu Trp Glu
 475 480 485 490

gag agg tgt ttt gtt tag ctctaga 1570
 Glu Arg Cys Phe Val
 495

<210> 2
 <211> 495
 45 <212> PRT
 <213> Cladorrhinum foecundissimum

<400> 2
 Met Leu Ile Leu Met Ile Pro Leu Phe Ser Tyr Leu Ala Ala Ala Ser
 50 1 5 10 15

Leu Arg Val Leu Ser Pro Gln Pro Val Ser Cys Asp Ser Pro Glu Leu
 20 25 30

55

Gly Tyr Gln Cys Asp Gln Gln Thr Thr His Thr Trp Gly Gln Tyr Ser
 35 40 45

Pro Phe Phe Ser Val Pro Ser Glu Ile Ser Pro Ser Val Pro Asp Gly
 50 55 60

5 Cys Arg Leu Thr Phe Ala Gln Val Leu Ser Arg His Gly Ala Arg Phe
 65 70 75 80

Pro Thr Pro Gly Lys Ala Ala Ala Ile Ser Ala Val Leu Thr Lys Ile
 85 90 95

10 Lys Thr Ser Ala Thr Trp Tyr Gly Ser Asp Phe Gln Phe Ile Lys Asn
 100 105 110

Tyr Asp Tyr Val Leu Gly Val Asp His Leu Thr Ala Phe Gly Glu Gln
 115 120 125

Glu Met Val Asn Ser Gly Ile Lys Phe Tyr Gln Arg Tyr Ser Ser Leu
 130 135 140

20 Ile Gln Thr Glu Asp Ser Asp Thr Leu Pro Phe Val Arg Ala Ser Gly
 145 150 155 160

Gln Glu Arg Val Ile Ala Ser Ala Glu Asn Phe Thr Thr Gly Phe Tyr
 165 170 175

25 Ser Ala Leu Ser Ala Asp Lys Asn Pro Pro Ser Ser Leu Pro Arg Pro
 180 185 190

Glu Met Val Ile Ile Ser Glu Glu Pro Thr Ala Asn Asn Thr Met His
 195 200 205

His Gly Leu Cys Arg Ser Phe Glu Asp Ser Thr Thr Gly Asp Gln Ala
 210 215 220

35 Gln Ala Glu Phe Ile Ala Ala Thr Phe Pro Pro Ile Thr Ala Arg Leu
 225 230 235 240

Asn Ala Gln Gly Phe Lys Gly Val Thr Leu Ser Asn Thr Asp Val Leu
 245 250 255

40 Ser Leu Met Asp Leu Cys Pro Phe Asp Thr Val Ala Tyr Pro Leu Ser
 260 265 270

Ser Leu Thr Thr Thr Ser Ser Val Ser Gly Gly Gly Lys Leu Ser Pro
 275 280 285

Phe Cys Ser Leu Phe Thr Ala Ser Asp Trp Thr Ile Tyr Asp Tyr Leu
 290 295 300

50 Gln Ser Leu Gly Lys Tyr Tyr Gly Phe Gly Pro Gly Asn Ser Leu Ala
 305 310 315 320

Ala Thr Gln Gly Val Gly Tyr Val Asn Glu Leu Ile Ala Arg Leu Ile
 325 330 335

55 Arg Ala Pro Val Val Asp His Thr Thr Thr Asn Ser Thr Leu Asp Gly

				340				345						350					
	Asp	Glu	Lys	Thr	Phe	Pro	Leu	Asn	Arg	Thr	Val	Tyr	Ala	Asp	Phe	Ser			
			355					360					365						
5	His	Asp	Asn	Asp	Met	Met	Asn	Ile	Leu	Thr	Ala	Leu	Arg	Ile	Phe	Glu			
		370					375					380							
	His	Ile	Ser	Pro	Met	Asp	Asn	Thr	Thr	Ile	Pro	Thr	Asn	Tyr	Gly	Gln			
10		385				390					395					400			
	Thr	Gly	Asp	Asp	Gly	Val	Lys	Glu	Arg	Asp	Leu	Phe	Lys	Val	Ser	Trp			
					405					410					415				
15	Ala	Val	Pro	Phe	Ala	Gly	Arg	Val	Tyr	Phe	Glu	Lys	Met	Val	Cys	Asp			
				420					425					430					
	Ala	Asp	Gly	Asp	Gly	Lys	Ile	Asp	Ser	Asp	Glu	Ala	Gln	Lys	Glu	Leu			
20			435					440					445						
	Val	Arg	Ile	Leu	Val	Asn	Asp	Arg	Val	Met	Arg	Leu	Asn	Gly	Cys	Asp			
		450					455					460							
	Ala	Asp	Glu	Gln	Gly	Arg	Cys	Gly	Leu	Glu	Lys	Phe	Val	Glu	Ser	Met			
25		465				470					475					480			
	Glu	Phe	Ala	Arg	Arg	Gly	Gly	Glu	Trp	Glu	Glu	Arg	Cys	Phe	Val				
					485					490					495				

PCT

5618-KaPe

Original (for SUBMISSION) - printed on 22.03.1999 10:06:53 AM


0-1	Form - PCT/RO/134 (EASY) Indications Relating to Deposited Microorganism(s) or Other Biological Material (PCT Rule 13bis)	
0-1-1	Prepared using	PCT-EASY Version 2.83 (updated 01.03.1999)
0-2	International Application No	
0-3	Applicant's or agent's file reference	5618-KaPe
1	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
1-1	page	6
1-2	line	18
1-3	Identification of Deposit	
1-3-1	Name of depositary institution	Centraalbureau voor Schimmelcultures
1-3-2	Address of depositary institution	Oosterstraat 1, Postbus 273, NL-3740 AG Baarn, Netherlands
1-3-3	Date of deposit	23 January 1997 (23.01.1997)
1-3-4	Accession Number	CBS 427.97
1-4	Additional Indications	NONE
1-5	Designated States for Which Indications are Made	all designated States
1-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE
2	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
2-1	page	6
2-2	line	20
2-3	Identification of Deposit	
2-3-1	Name of depositary institution	DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
2-3-2	Address of depositary institution	Mascheroder Weg 1b, D-38124 Braunschweig, Germany
2-3-3	Date of deposit	17 March 1999 (17.03.1999)
2-3-4	Accession Number	DSMZ 12742
2-4	Additional Indications	NONE
2-5	Designated States for Which Indications are Made	all designated States
2-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

PCT

5618-KaPe

Original (for SUBMISSION) - printed on 22.03.1999 10:06:53 AM

FOR RECEIVING OFFICE USE ONLY

0-4	This form was received with the international application: (yes or no)	yes
0-4-1	Authorized officer	

FOR INTERNATIONAL BUREAU USE ONLY

0-5	This form was received by the international Bureau on:	
0-5-1	Authorized officer	

INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 99/00153

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: C12N 9/16, C12N 15/55, A23K 1/165
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: C12N, A23K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	EP 0897010 A2 (F. HOFFMANN-LA ROCHE AG), 7 February 1999 (07.02.99) --	1-47
P,X	EP 0897985 A2 (F. HOFFMANN-LA ROCHE AG), 24 February 1999 (24.02.99) --	1-47
X	WO 9735016 A1 (NOVO NORDISK BIOTECH, INC.), 25 Sept 1997 (25.09.97), page 10, line 22 - page 11, line 18, claim 11 --	1-47
X	EP 0420358 A1 (GIST-BROCADES N.V.), 3 April 1991 (03.04.91), page 10, line 6 - line 14; and the claims --	1-47

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

6 July 1999

Date of mailing of the international search report

13 -07- 1999

Name and mailing address of the ISA:

Swedish Patent Office
Box 5055, S-102 42 STOCKHOLM

Authorized officer

Carolina Palmcrantz/Els

INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 99/00153

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 9748812 A2 (HER MAJESTY THE QUEEN IN RIGHT OF CANADA), 24 December 1997 (24.12.97), page 13, line 7 - line 12 --	1-47
A	WO 9114782 A1 (GIST-BROCADES N.V.), 3 October 1991 (03.10.91) -- -----	1-47

INTERNATIONAL SEARCH REPORT
Information on patent family members

01/06/99

International application No.
PCT/DK 99/00153

Patent document cited in search report			Publication date	Patent family member(s)		Publication date
EP	0897010	A2	07/02/99	AU	5954398 A	01/10/98
				CA	2231948 A	25/09/98
				JP	10276789 A	20/10/98
EP	0897985	A2	24/02/99	NO	983364 D	00/00/00
WO	9735016	A1	25/09/97	AU	2077197 A	10/10/97
				AU	2539197 A	10/10/97
				CA	2248980 A	25/09/97
				EP	0904383 A	31/03/99
				US	5866118 A	02/02/99
				WO	9735017 A	25/09/97
EP	0420358	A1	03/04/91	AT	180014 T	15/05/99
				AU	636673 B	06/05/93
				AU	6501190 A	28/04/91
				BG	60108 A	15/10/93
				CA	2042054 A	28/03/91
				CN	1051058 A	01/05/91
				DE	420358 T	12/10/95
				EP	0779037 A	18/06/97
				ES	2072834 T	01/08/95
				FI	912530 D	00/00/00
				HU	215179 B	28/10/98
				JP	4506007 T	22/10/92
				LT	1527 A	26/06/95
				LT	3957 B	27/05/96
				LV	10310 A,B	20/10/94
				NO	303988 B	05/10/98
				PL	167790 B	30/11/95
				PL	168470 B	29/02/96
				PT	95447 A,B	22/05/91
				US	5436156 A	25/07/95
				US	5863533 A	26/01/99
				WO	9105053 A	18/04/91
WO	9748812	A2	24/12/97	AU	3021697 A	07/01/98
				CA	2257101 A	24/12/97
				EP	0904385 A	31/03/99
				NO	985804 D	00/00/00

INTERNATIONAL SEARCH REPORT
Information on patent family members

01/06/99

International application No.
PCT/DK 99/00153

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9114782 A1	03/10/91	AU 632941 B	14/01/93
		AU 649447 B	26/05/94
		AU 7765691 A	21/10/91
		AU 7776691 A	21/10/91
		CA 2054762 A	24/09/91
		CA 2056396 A	24/09/91
		EP 0449375 A	02/10/91
		EP 0449376 A	02/10/91
		FI 915477 D	00/00/00
		FI 915478 D	00/00/00
		HU 215164 B	28/10/98
		HU 215260 B	30/11/98
		IL 97645 A	18/03/97
		JP 6501838 T	03/03/94
		JP 6502296 T	17/03/94
		PT 97110 A,B	29/11/91
		PT 97111 A	31/12/91
		US 5543576 A	06/08/96
		US 5714474 A	03/02/98
		WO 9114772 A	03/10/91
		US 5593963 A	14/01/97
		US 5770413 A	23/06/98